

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 27.2165 Seconds
(without alignments)
46.656 Million cell updates/sec

Title: US-09-991-433-6
Perfect score: 44
Sequence: 1 NKGTQOXYT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	370	AA13406	Parvo virus B19 PA
2	44	100.0	543	AA13405	Parvo virus B19 VP
3	44	100.0	554	AAW08987	Human parvovirus V
4	44	100.0	554	AA23230	Erythrovirus V9 VP
5	44	100.0	554	ABP57264	Human parvovirus B
6	44	100.0	554	ABP57267	Human parvovirus B
7	44	100.0	756	AA171231	Capsid protein enc
8	44	100.0	781	AAW08986	Human parvovirus V
9	44	100.0	781	AA23227	Erythrovirus V9 VP

10	44	100.0	781	24	ABP57263	Human parvovirus B
11	44	100.0	781	24	ABP57266	Human parvovirus B
12	35	79.5	237	22	ABW60093	Drosophila melanog
13	33	75.0	112	23	ABP32844	Human ORF1817 prot
14	32	72.7	277	22	ABW64125	Drosophila melanog
15	32	72.7	326	22	AAW39794	Human polypeptide
16	32	72.7	326	22	AAW39794	Human polypeptide
17	32	72.7	365	22	AAW39794	Human polypeptide
18	32	72.7	365	22	AAW39794	Human polypeptide
19	31	70.5	239	23	ABW78123	Amino acid sequenc
20	31	70.5	239	23	ABW78123	Amino acid sequenc
21	31	70.5	260	22	AAW79123	Partial elastase s
22	31	70.5	667	14	AAW39573	Sequence of 741 SF
23	31	70.5	727	22	AAW39573	S. epidermidis ope
24	31	70.5	738	23	ABP39440	Staphylococcus epi
25	31	70.5	1037	22	ABW04999	Novel human diagno
26	31	70.5	1837	22	ABW62913	Drosophila melanog
27	30	68.2	13	21	AAW51863	Pancreatic elastas
28	30	68.2	98	19	AAW40575	Human IG52 CH1 dom
29	30	68.2	98	20	AAW08742	Human IG52 protein
30	30	68.2	98	20	AAW2422	Human IG52 protein
31	30	68.2	98	21	AAW26997	Human IG52 CH1 dom
32	30	68.2	98	23	ABW31880	Human IG52 CH1 dom
33	30	68.2	116	18	AAW41393	Human heavy chain
34	30	68.2	116	20	AAW5391	Chimaeric antibody
35	30	68.2	235	14	AAW41707	Murine 128.1 VH/hu
36	30	68.2	240	8	AAW70045	Human pancreatic e
37	30	68.2	250	17	AAW06179	Humanised A5B57 Pd
38	30	68.2	252	8	AAW70047	Human chromosomal
39	30	68.2	255	18	AAW41394	Chimeric anti-CEA
40	30	68.2	255	18	AAW41399	Humanised antibody
41	30	68.2	258	21	AAW30513	Amino acid sequenc
42	30	68.2	258	23	ABW43190	Human ovarian anti
43	30	68.2	266	6	AAW50735	Porcine pancreatic
44	30	68.2	266	6	AAW50360	Pig elastase I gen
45	30	68.2	266	8	AAW70757	Pig pancreas elast

ALIGNMENTS

RESULT 1	AA13406	AA13406 standard; Protein; 370 AA.
ID	AA13406	
AC	AA13406	
DT	24-OCT-1991	(first entry)
DE	Parvo virus B19 PANSE.	
KW	Primer; PCR; PAPST; globulin.	
OS	Synthetic.	
PN	DE4003826-A.	
PD	14-AUG-1991.	
PF	08-FEB-1990;	90DE-4003826.
PR	08-FEB-1990;	90DE-4003826.
PA	(MIR-) MIKROGEN MOLEKULARB.	
PI	Soutschek E, Motz M;	
DR	WPI, 1991-246423/34.	
XX	Immunologically active parvo virus B19 peptide(s) - comprising	
PT	capsid protein VPI or VP2 fragments, useful for antibody	
XX	detection or vaccination	

PS Claim 16; Page 10-11; 22pp; German.
 XX
 CC VP2 (AA13405) and its fragments PANSE and PAPST (AA13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin prepns.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AA013159-63 in pairs.
 CC See also AA13400-07 and AA13414.
 CC
 XX
 SQ Sequence 370 AA;

Query Match 100.0%; Score 44; DB 12; Length 370;
 Best Local Similarity 100.0%; Pred. No. 0.76;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYT 8
 |||||
 231 NKGTQOYT 238

RESULT 2
 AA13405
 ID AA13405 standard; Protein; 543 AA.

AA13405;

24-OCT-1991 (first entry)

Parvo virus B19 VP2.

Primer; PCR; globulin; PANSE; PAPST.

Synthetic.

DE4003826-A.

14-AUG-1991.

08-FEB-1990; 90DE-4003826.

08-FEB-1990; 90DE-4003826.

(MIKR-) MIKROGEN MOLEKULARB.

Soutschek E, Mottz M;

WPI; 1991-246423/34.

Immunologically active parvo virus B19 peptide(s) - comprising
 capsid protein VP1 or VP2 fragments, useful for antibody
 detection or vaccination

Diclosure; Fig 2-6; 22pp; German.

VP2 and its fragments PANSE (AA13406) and PAPST (AA13407) are useful
 as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 for diagnosis of B19 infections, determining the immune status of
 pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin prepns.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AA013159-63 in pairs.
 CC See also AA13400-07 and AA13414.
 CC
 XX
 SQ Sequence 543 AA;

Query Match 100.0%; Score 44; DB 12; Length 543;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYT 8
 |||||
 Db 404 NKGTQOYT 411

RESULT 3
 AA08987
 ID AA08987 standard; Protein; 554 AA.

AA08987;

27-FEB-1997 (first entry)

Human parvovirus VP-2 protein.

Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 erythblastemia; abortion; universal fetal hydrops; liver disease;
 haemorrhagic fever; rheumatism; detection; IgG antibody.

Human parvovirus.

JP07147986-A.

13-JUN-1995.

24-SEP-1992; 92JP-0281017.

24-SEP-1992; 92JP-0281017.

(DENK-) DENKA SEIKEN KK.

(ELED) DENKI KAGAKU KOGYO KK.

WPI; 1995-242756/32.

N-PSDB; AAT49535.

Human parvovirus gene coding for a polypeptide - useful for
 developing vaccines against parvoviral diseases such as
 erythroblastemia, haemorrhagic fever, etc.

Claim 3; Page 7-9; 38pp; English.

The sequences given in AA08986 represent the parvovirus structural
 proteins, VP-1 and VP2, and the non-structural protein, NS. The
 genomic fragment encoding these proteins may be used for the
 diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.

Sequence 554 AA;

Query Match 100.0%; Score 44; DB 16; Length 554;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYT 8
 |||||
 Db 395 NKGTQOYT 402

RESULT 4
 AA023230
 ID AA023230 standard; Protein; 554 AA.

AA023230;

26-AUG-1999 (first entry)

Erythrovirus V9 VP2 protein.

Erythrovirus V9; differential diagnosis; parvovirus; infection;
 erythrovirus screening; typing; immunoassay; VP2 protein.

```

XX OS Erythrovirus.
XX PN FR2771751-A1.
XX PD 04-JUN-1999.
XX PF 03-DEC-1997; 97FR-0015197.
XX PR 03-DEC-1997; 97FR-0015197.
XX PA (ASSI-) ASSISTANCE PUBLIQUE HOPITALUX PARIS.
XX PI Auguste V, Garbary CA, Nguyen QT;
XX DR MPI; 1999-349543/30.
XX DR N-PSDB; AAX81586.
XX PT Erythrovirus V9 and its nucleic acid sequences - can be used in the
    diagnosis of its infections
    Claim 19; Page 57-58; 80pp; French.
XX CC The present sequence represents an erythrovirus V9 protein.
XX CC Probes and primers derived from erythrovirus V9 polynucleotide
XX CC sequences (AAX81580) can be used for differential diagnosis of
XX CC erythrovirus (parvovirus) infections by a combination of
XX CC amplification and hybridisation assay. The probes can also be
XX CC used to assess susceptibility to erythrovirus infection and
XX CC for erythrovirus screening and typing. The antibodies can be
XX CC used in immunoassays for diagnosis of erythrovirus V9 infections.
XX SQ Sequence 554 AA;
    Query Match 100.0%; Score 44; DB 20; Length 554;
    Best Local Similarity 100.0%; Pred. No. 1.2;
    Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    Oy 1 NKGTQOY 8
    Db 395 NKGTQOY 402
RESULT 5
ABP57264
ID ABP57264 standard; Protein; 554 AA.
XX ABP57264;
    22-APR-2003 (first entry)
XX DE Human parvovirus B19 clone B1-VP2 amino acid sequence SEQ ID NO:29.
XX KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
XX OS Human parvovirus B19.
XX PN WO2003002753-A2.
XX PD 09-JAN-2003.
XX PF 28-JUN-2002; 2002WO-US20684.
XX PR 28-JUN-2001; 2001US-302077P.
XX PR 19-MAR-2002; 2002US-365956P.
XX PR 29-MAR-2002; 2002US-369224P.
XX PA (CHIR ) CHIRON CORP.
XX PI Pichuanes S, Shyamala V;
XX DR MPI; 2003-201510/19.
XX DR N-PSDB; AB259574.

```

```

XX PT Detecting a human parvovirus B19 infection in a biological sample to
XX PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
XX PT acid with a primer complementary to the 3'-terminal portion of the RNA
XX PT target sequence
XX PS Example 4; Fig 7B; 148pp; English.
XX CC The present invention describes a method for detecting a human parvovirus
XX CC B19 infection in a biological sample. The method comprises reacting the
XX CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
XX CC consisting of a first primer containing a complexing sequence
XX CC sufficiently complementary to the 3'-terminal portion of the RNA target
XX CC sequence to complex with. Also described: (1) amplifying a target
XX CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
XX CC of 47 700 base pair sequences (see AB259549 to AB259569, and AB259604 to
XX CC AB259629); (3) a polynucleotide comprising either of 2 4678 base pair
XX CC sequences (see AB259570 and AB259571); (4) an oligonucleotide primer
XX CC consisting of a promoter region recognised by a DNA-dependent RNA
XX CC polymerase operably linked to a human parvovirus B19-specific complexing
XX CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
XX CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
XX CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
XX CC oligonucleotide primer of (4), and instructions for conducting the
XX CC diagnostic test. The method is useful for detecting parvovirus infection
XX CC in a biological sample, such as in blood products, to prevent
XX CC transmission of the virus through blood and plasma derivatives or by
XX CC close personal contact. AB259549 to AB259634 and ABP57262 to ABP57267
XX CC represent sequences used in the exemplification of the present invention.
XX SQ Sequence 554 AA;
    Query Match 100.0%; Score 44; DB 24; Length 554;
    Best Local Similarity 100.0%; Pred. No. 1.2;
    Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    Oy 1 NKGTQOY 8
    Db 395 NKGTQOY 402
RESULT 6
ABP57267
ID ABP57267 standard; Protein; 554 AA.
XX ABP57267;
    22-APR-2003 (first entry)
XX DE Human parvovirus B19 clone B6-VP2 amino acid sequence SEQ ID NO:35.
XX KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
XX OS Human parvovirus B19.
XX PN WO2003002753-A2.
XX PD 09-JAN-2003.
XX PF 28-JUN-2002; 2002WO-US20684.
XX PR 28-JUN-2001; 2001US-302077P.
XX PR 19-MAR-2002; 2002US-365956P.
XX PR 29-MAR-2002; 2002US-369224P.
XX PA (CHIR ) CHIRON CORP.
XX PI Pichuanes S, Shyamala V;
XX DR MPI; 2003-201510/19.
XX DR N-PSDB; AB259577.
XX PT Detecting a human parvovirus B19 infection in a biological sample to

```

PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic acid with a primer complementary to the 3'-terminal portion of the RNA target sequence -

PS Example 4; Fig 10B; 148pp; English.

XX The present invention describes a method for detecting a human parvovirus B19 infection in a biological sample. The method comprises reacting the isolated parvovirus B19 nucleic acid with a first oligonucleotide consisting of a first primer containing a complexing sequence sufficiently complementary to the 3'-terminal portion of the RNA target sequence to complex with. Also described: (1) amplifying a target parvovirus B19 nucleic acid sequence; (2) a polynucleotide comprising one of 47 700 base pair sequences (see AB259549 to AB259569, and AB259604 to AB259629); (3) a polynucleotide comprising either of 2 4678 base pair sequences (see AB259570 and AB259571); (4) an oligonucleotide primer consisting of a promoter region recognised by a DNA-dependent RNA polymerase operably linked to a human parvovirus B19-specific complexing sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked to an acridinium ester label; and (6) a diagnostic test kit comprising an oligonucleotide primer of (4), and instructions for conducting the diagnostic test. The method is useful for detecting parvovirus infection in a biological sample, such as in blood products, to prevent transmission of the virus through blood and plasma derivatives or by close personal contact. AB259549 to AB259634 and ABP57262 to ABP57267 represent sequences used in the exemplification of the present invention.

CC Sequence 554 AA;

CC Query Match 100.0%; Score 44; DB 24; Length 554;

CC Best Local Similarity 100.0%; Pred. No. 1.2;

CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8

DB 395 NKGTOOYT 402

RESULT 7

ID AAV71231 standard; Protein; 756 AA.

XX AAV71231;

XX 08-SEP-2000 (first entry)

XX Capsid protein encoded by AAV2/human parvovirus B19 chimeric vector.

XX Adeno-associated virus; AAV2; human parvovirus B19; chimeric; recombinant parvoviral vector; cellular tropism; cap protein; capsid; gene delivery; gene therapy; VP1; VP2; VP3.

XX Chimeric - Adeno associated virus serotype 2.

XX Chimeric - Human parvovirus B19.

XX WO20028004-A1.

XX 18-MAY-2000.

XX 10-NOV-1999; 99WO-US26505.

XX 10-NOV-1998; 98US-0107840.

XX 10-MAR-1999; 99US-0123651.

XX (UNCL) UNIV NORTH CAROLINA.

XX Rabinowitz JE, Samulski RJ, Xiao W;

XX MPI: 2000-376523/32.

XX N-PSDB; AAD00833.

XX Recombinant parvoviral vectors with altered packaging, tropisms and

PT immunogenic properties, useful in gene therapy protocols -

XX Example 21; Page 142; 153pp; English.

XX The patent discloses modified parvovirus vectors with advantageous antigenic properties, packaging capabilities and cellular tropisms. These vectors can be used in standard recombinant DNA protocols e.g. gene therapy for delivering nucleic acids to cells.

XX The present sequence is a capsid protein encoded by an adeno-associated virus serotype 2 (AAV2)/human parvovirus B19 chimeric vector. This vector encodes AAV2 VP1 and VP2 capsid proteins and human parvovirus B19 VP2 protein. The chimeric vector was constructed by replacing the VP3 major cap protein of AAV2 with B19's VP2. Recombinant parvovirus comprising the chimeric capsid is useful for gene delivery.

CC Sequence 756 AA;

CC Query Match 100.0%; Score 44; DB 21; Length 756;

CC Best Local Similarity 100.0%; Pred. No. 1.7;

CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8

DB 597 NKGTOOYT 604

RESULT 8

ID AAM08986 standard; Protein; 781 AA.

XX AAM08986;

XX 27-FEB-1997 (first entry)

XX Human parvovirus VP-1 protein.

XX Human parvovirus genome; structural gene; VP-1; VP2; arthritis; non-structural protein; NS; diagnosis; vaccine; parvoviral disease; erythraemia; abortion; universal fetal hydrops; liver disease; haemorrhagic fever; rheumatism; detection; IgG antibody.

XX Human parvovirus.

XX Key Location/Qualifiers

XX Misc-difference 19 /label= Gly, Val

XX Misc-difference 61 /label= Asn, Asp

XX Misc-difference 220 /label= His, Asn

XX JP07147986-A.

XX 13-JUN-1995.

XX 24-SEP-1992; 92JP-0281017.

XX 24-SEP-1992; 92JP-0281017.

XX (DENK-) DENKA SEIKEN KK.

XX (ELED) DENKI KAGAKU KOGYO KK.

XX MPI: 1995-242756/32.

XX N-PSDB; AAT49535.

XX Human parvovirus gene coding for a polypeptide - useful for developing vaccines against parvoviral diseases such as erythroblastemia, haemorrhagic fever, etc.

XX Claim 2; Page 5-7; 38pp; English.

XX The sequences given in AAM08986 represent the parvovirus structural

CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythraemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.
 SQ Sequence 781 AA;

Query Match 100.0%; Score 44; DB 16; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTOOYT 8
 |||||
 Db 622 NKGTOOYT 629

RESULT 9
 3227
 AAY23227 standard; Protein; 781 AA.

AC AAY23227;
 DT 26-AUG-1999 (first entry)
 XX
 XX Erythrovirus V9 VP1 protein.

DE Erythrovirus V9 VP1 protein.
 XX
 XX Erythrovirus V9; differential diagnosis; parvovirus; infection;
 KM erythrovirus screening; typing; immunoassay; VP1 protein.

OS Erythrovirus.

PN FR2771751-A1.

PD 04-JUN-1999.

PF 03-DEC-1997; 97FR-0015197.

PR 03-DEC-1997; 97FR-0015197.

XX (ASST-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

PI Auguste V, Garbarg CA, Nguyen QT;

XX WPI; 1999-349543/30.

DR N-PSDB; AAX81583.

Erythrovirus V9 and its nucleic acid sequences - can be used in the
 diagnosis of its infections

PS Claim 19; Page 50-52; 80pp; French.

XX The present sequence represents an erythrovirus V9 protein.
 CC Probes and primers derived from erythrovirus V9 polynucleotide
 CC sequences (AAX81580) can be used for differential diagnosis of
 CC erythrovirus (parvovirus) infections by a combination of
 CC amplification and hybridisation assay. The probes can also be
 CC used to assess susceptibility to erythrovirus infection and
 CC for erythrovirus screening and typing. The antibodies can be
 CC used in immunoassays for diagnosis of erythrovirus V9 infections.
 XX

SQ Sequence 781 AA;

Query Match 100.0%; Score 44; DB 20; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTOOYT 8
 |||||
 Db 622 NKGTOOYT 629

RESULT 10

ABP57263
 ID ABP57263 standard; Protein; 781 AA.

XX ABP57263;

DT 22-APR-2003 (first entry)

DE Human parvovirus B19 clone B1-VP1 amino acid sequence SEQ ID NO:27.

KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.

XX Human parvovirus B19.

OS WO2003002753-A2.

PN 09-JAN-2003.

PF 28-JUN-2002; 2002WO-US20684.

PR 28-JUN-2001; 2001US-302077P.

PR 19-MAR-2002; 2002US-365956P.

PR 29-MAR-2002; 2002US-369224P.

XX (CHIR) CHIRON CORP.

PI Pichuanes S, Shyamala V;

XX WPI; 2003-201510/19.

DR N-PSDB; AB259573.

PT Detecting a human parvovirus B19 infection in a biological sample to
 PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -

XX Example 4; Fig 6B; 148pp; English.

PS The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see AB259549 to AB259569, and AB259604 to
 CC AB259629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see AB259570 and AB259571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. AB259549 to AB259574 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.
 XX

SQ Sequence 781 AA;

Query Match 100.0%; Score 44; DB 24; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTOOYT 8
 |||||
 Db 622 NKGTOOYT 629

RESULT 11
 ABP57266

ID	ABP57266 standard; Protein; 761 AA.
XX	
AC	ABP57266;
XX	
DT	22-APR-2003 (first entry)
XX	
DE	Human parvovirus B19 clone B6-VPI amino acid sequence SEQ ID NO:33.
XX	
KV	Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
XX	
OS	Human parvovirus B19.
XX	
PN	WO2003002753-A2.
XX	
PD	09-JAN-2003.
XX	
PF	28-JUN-2002; 2002WO-US20684.
XX	
PR	28-JUN-2001; 2001US-302077P. 19-MAR-2002; 2002US-365956P. 29-MAR-2002; 2002US-369224P.

PA (CHIR) CHIRON CORP.
XX
PI Pichuanetes S, Shyamala V;
XX
DR MPI; 2003-201510/19.
DR N-PsDB; ABZ59576.
XX
XX Detecting a human parvovirus B19 infection in a biological sample to
PT prevent viral transamination, comprises reacting a parvovirus B19 nucleic
PT acid with a primer complementary to the 3'-terminal portion of the RNA
PT target sequence -

Example 4; Fig 9B; 148pp; English.

CC The present invention describes a method for detecting a human parvovirus
CC B19 infection in a biological sample. The method comprises reacting the
CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
CC consisting of a first primer containing a complexing sequence
CC sufficiently complementary to the 3'-terminal portion of the RNA target
CC sequence to complex with. Also described: (1) amplifying a target
CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
CC of 47 700 base pair sequences (see ABZ59549 to ABZ55569, and ABZ59604 to
CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
CC consisting of a promoter region recognised by a DNA-dependent RNA
CC polymerase operably linked to a human parvovirus B19-specific complexing
CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
CC to an acidindium ester label; and (6) a diagnostic test kit comprising an
CC oligonucleotide primer of (4), and instructions for conducting the
CC diagnostic test. The method is useful for detecting parvovirus infection
CC in a biological sample, such as in blood products, to prevent
CC transmission of the virus through blood and plasma derivatives or by
CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
CC represent sequences used in the exemplification of the present invention.
XQ
XQ Sequence 781 Aa;

Query Match	100.0%	Score 44;	DB 24;	length 781;
Best Local Similarity	100.0%;	Pred. NO. 1.7;		
Matches	8;	Conservative	0;	Indels 0;
		Mismatches	0;	Gaps 0;

QY 1 NKGTQQT 8
Db 622 NKGTQQT 629

RESULT 12
ABB60093
ID ABB60093 standard; Protein; 237 AA.

AC	ABB60093;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 7071.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
PD	
XX	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW,
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL04196.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 7071; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 237 AA;

Query Match	79.5%	Score 35;	DB 22;	length 237;
Best Local Similarity	75.0%	Pred. NC. 34;		
Matches	6;	Conservative	1;	Mismatches 0;
			Indels	Gaps 0

QY	1	8
	:	:
Db	30	37
	NNGTQKYT	

RESULT 13
ABP32844
ID ABP32844 standard; Protein; 112 AA

AC	ABP32844;
XX	
DT	09-JUL-2002 (first entry)

DE	Human ORF1817 protein, SEQ ID NO:3634.
XX	
XX	
KW	Human; ORF; open reading frame; ORFX;

KM Human; ORP: open reading frame; ORP: drug screening; diagnosis;
KM disease monitoring; cytokine; cell proliferation; cell differentiation;
KM immune modulation; haematopoiesis regulation; tissue growth
KM angiogenesis; activin; inhibitor; chemotactic; chemokinetic; haemostatic;
KM thrombolytic; tumour inhibition; biologically characteristic; fertility;
KM behavioural cancer; proliferative disorder; neurological disorder;
KM cardiovascular disease; immune system disorder; organ transplantation;

KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnary;
 KW vasotropic; antipsoriatic; antidiabetic; cytosolic; nocrotic;
 KW neuroprotective; antihypertensive; anticoagulant; thrombolytic;
 KW cardiatic; hypotensive; antithyroid; antinflammatory; immunomodulator;
 KW dermatological; analgesic; vitruicide; antibacterial; fungicide.
 XX Homo sapiens.
 OS
 XX MO200190366-A2.
 PN
 XX 29-NOV-2001.
 PD
 XX 24-MAY-2001; 2001WO-US17076.
 PF
 XX 24-MAY-2000; 2000US-206690P.
 PR
 XX (CUBA-) CUBAGEN CORP.
 PA
 XX Leach MD, Shinkels RA;
 XX
 DR MPI; 2002-106200/14.
 N-PSDB; ABN76870.
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 XX
 XX Claim 10; Page 1167; 2508pp; English.
 PS
 XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, tumour inhibitory activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 CC
 XX
 SQ Sequence 112 AA;
 Query Match 75.0%; Score 33; DB 23; Length 112;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 KGTQOYT 8

Db |||||
 21 KGAQOYT 27
 RESULT 14
 ID ABB64125 standard; Protein; 277 AA.
 XX ABB64125;
 AC ABB64125;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 19167.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS
 XX MO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX
 PI MPI; 2001-656860/75.
 DR N-PSDB; ABL08228.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Disclosure; SEQ ID NO 19167; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS57737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 277 AA;
 Query Match 72.7%; Score 32; DB 22; Length 277;
 Best Local Similarity 71.4%; Pred. No. 1.7e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NKGTQOYT 7
 Db 247 NRGTRQY 253
 RESULT 15
 ID AAM39794 standard; Protein; 326 AA.
 XX AAM39794;
 AC AAM39794;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2939.

XX	Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer
KM	peripheral nervous system; neuropathy; central nervous system; CNS;
KM	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM	leukaemia.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US4263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-058047.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693096.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.

Query Match	72.7%	Score 32;	DB 22;	Length 326;
Best Local Similarity	62.5%	Pred. No. 26+02;		
Matches	5; Conservative	2; Mismatches	1; Indels	0; Gaps
OY	1 NKGITQY ^T 8			
db	249 NKGITHYS 256			

RESULT 16	
AAB92876	
ID	AAB92876 standard; Protein; 326 AA
XX	
AC	AAB92876;

DT	26-JUN-2001	(first entry)
XX		
DE	Human protein sequence SEQ ID NO:11468.	
XX		
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
PN	EP1074617-A2.	
XX		
PD	07-FEB-2001.	
XX		
PE	28-JUL-2000; 2000EP-0116126.	
XX		
PR	29-JUL-1999; 99JP-0248036.	
PR	27-AUG-1999; 99JP-0300253.	
PR	11-JAN-2000; 2000JP-0118776.	
PR	02-MAY-2000; 2000JP-0183767.	
PR	09-JUN-2000; 2000JP-0241899.	
XX		
PA	(HELI-) HELIX RES INST.	
XX		
PI	Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J,	
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
XX		
XX	WPI; 2001-318749/34.	

	Query Match	72.7%	Score 32;	DB 22;	length 326;
	Best Local Similarity	62.5%	Pred. No.	2e+02;	
	Matches	5;	Conservative	2;	Mismatches 1;
					Indels 0; Gaps 0;
Oy	1 NKGTOOFT 8				
	:				
Db	249 NKGTENY 256				

RESULT 17
AAG73503
ID AAG73503 standard; Protein; 365 AA
XX
AC AAG73503;

XX 10-AUG-2001 (first entry)
DE Human gene 30-encoded secreted protein fragment, SEQ ID NO:279.
XX
XX Human, secreted protein; proliferative disorder; cancer;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiotensin disorder; kidney disorder;
KW gastroenteric disorder; pregnancy-related disorder; tumour;
KW endocrine disorder; infection; wound healing; vulnery;
KW cell culture; chemotaxis; food additive;
KW binding partner identification.
XX Homo sapiens.
OS
XX WO200134628-A1.
XX 17-MAY-2001.
XX 08-NOV-2000; 2000WO-US30653.
XX 12-NOV-1999; 99US-0164735.
XX 27-JUL-2000; 2000US-0221193.
XX (HOMA-) HUMAN GENOME SCT INC.
XX Ruben SM, Komatoulis GA, Birse CE, Ni J, Moore PA;
XX WPI; 2001-32906/34.
XX Nucleic acids encoding 35 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
XX
XX Disclosure; Page 66; 604pp; English.
XX AAH32522-AAH32627 represent cDNAs corresponding to 35 human secreted
XX protein genes, and AAG73449-AAG73448 represent the proteins they encode.
XX AAG73449-AAG73519 represent human secreted protein fragments. The genes
XX and their corresponding secreted proteins are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 52 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of
XX proliferative disorders, cancer, tumours, foetal and developmental
XX abnormalities, haematopoietic disorders, diseases of the immune system,
XX AIDS, autoimmune diseases (e.g., Rheumatoid arthritis), inflammation,
XX allergies, neurological disorders (e.g., Alzheimer's disease,
XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,
XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
XX cardiovascular disorders, angiogenic disorders, kidney disorders,
XX gastrointestinal disorders, pregnancy-related disorders, endocrine
XX disorders, and infections. The proteins can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues, to identify their
XX cognate ligands or binding partners, and in chemotaxis, and can be used
XX as a food additive or preservative to modify storage properties.
XX Antibodies specific for a protein of the invention can be used in
XX alleviating symptoms associated with the disorders mentioned above, and
XX in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
XX immunoassay (ELISA). The present sequence represents a human
XX secreted protein fragment referred to in the disclosure of the invention.
XX
XX Sequence 365 AA;
XX

Query Match 72.7% Score 32; DB 22; Length 365;

Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CY 1 NKGTQVPT 8
DB 249 NKGTBHY 256
RESULT 18
AAM41580
ID AAM41580 standard; Protein, 371 AA.
XX
XX AAM41580;
AC
XX
XX 22-OCT-2001 (first entry)
DE Human polypeptide SEQ ID NO 6511.
XX
XX Human, noctropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
OS
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0558042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX (HYSB-) HYSBQ INC.
XX
XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao Q, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX N-PSDB; AAI60736.
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6511; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with noctropic,
XX immunosuppressant and cyostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX

SQ Sequence 371 AA;

Query Match 72.7%; Score 32; DB 22; Length 371;
 Best Local Similarity 62.5%; Pred. No. 2.3e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
 |||||:
 DB 294 NKGTHTYS 301

RESULT 19

AB878123
 ID AB878123 standard; peptide; 239 AA.

AC AB878123;

DT 05-NOV-2002 (first entry)

XX Amino acid sequence of elastase.

XX Atomic group sequence; protein; trypsin.

OS Unidentified.

PN US2002072863-A1.

PD 13-JUN-2002.

XX 23-JUL-2001; 2001US-0909809.

XX 06-FEB-1992; 92JP-0021012.

XX 11-DEC-1992; 92JP-0331703.

XX 08-FEB-1993; 93US-0014867.

XX (FUIT) FUJITSU LTD.

XX Tomikawa M, Aikawa S, Matsuzawa F;

XX WPI; 2002-607266/65.

Analysis of sequences of atomic groups for automatically extracting and
 evaluating mutually coinciding or similar portions between amino acid
 sequences in proteins, comprises obtaining a longest common atomic
 group number

PS Disclosure; Fig 38B; 65pp; English.

The specification describes a method for the analysis of sequences of
 atomic groups for automatically extracting and evaluating mutually
 coinciding or similar portions between amino acid sequences in proteins.
 The method comprises preparing an array S(1) having array elements S(0)
 to S (m); initializing all array elements of the array S(1) to 0 and
 initializing an integer j to 1; adding 1 to each array element and 1 to
 the integer j until j exceeds n; and obtaining a longest common atomic
 group number. The method is used for analysing sequences of atomic
 groups, particularly for automatically extracting and evaluating mutually
 coinciding or similar portions between amino acid sequences in protein
 molecules and/or between three-dimensional structures of protein
 molecules. The present sequence represents elastase, and is used
 in the course of the invention.

SQ Sequence 239 AA;

Query Match 70.5%; Score 31; DB 23; Length 239;
 Best Local Similarity 71.4%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
 |||||:
 DB 65 NKGTQOY 71

RESULT 20

AB83323
 ID AB83323 standard; protein; 239 AA.

AC AB83323;

DT 30-AUG-2002 (first entry)

XX Partial elastase sequence.

XX Elastase; three dimensional structure; protein structure analysis;

KW serine protease.

XX Unidentified.

XX US2002035434-A1.

XX 21-MAR-2002.

XX 23-JUL-2001; 2001US-0910054.

XX 06-FEB-1992; 92JP-0021012.

XX 11-DEC-1992; 92JP-0331703.

XX 08-FEB-1993; 93US-0014867.

XX (FUIT) FUJITSU LTD.

XX Tomikawa M, Aikawa S, Matsuzawa F;

XX WPI; 2002-507172/54.

Analysis of three-dimensional structures by generating combination of
 correspondence satisfying restriction condition, and calculating root
 mean square distance between elements in the combination of
 correspondence

PS Disclosure; Fig 38; 65pp; English.

The present invention relates to a method for analysis of three
 dimensional structures. The method involves generating a combination of
 correspondence satisfying a restriction condition between the elements
 belonging to a first and second point sets from among all candidates for
 the combination of correspondence, and calculating a root mean square
 distance between the elements corresponding in the combination of
 correspondence. The method is useful for analysing three-dimensional
 structures of molecules, particularly proteins. To illustrate the
 invention, the serine proteases trypsin and elastase were used. The
 present sequence is a partial sequence for elastase. The histidine and
 serine active sites of trypsin and elastase were analysed using the
 method of the invention.

SQ Sequence 239 AA;

Query Match 70.5%; Score 31; DB 23; Length 239;
 Best Local Similarity 71.4%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
 |||||:
 DB 65 NKGTQOY 71

RESULT 21

AA679123
 ID AA679123 standard; Protein; 260 AA.

AC AA679123;

DT 03-JAN-2002 (first entry)

XX Amino acid sequence comprising a phage resistance mechanism.

XX Phage resistance mechanism; lactic acid bacteria; cheese; yogurt;

KW butter; sauerkraut.
 XX Streptococcus thermophilus.
 XX MO200170990-A1.
 XX 27-SEP-2001.
 XX 21-MAR-2001; 2001WO-FR00836.
 XX 21-MAR-2000; 2000FR-0003578.
 XX (SUDD) SKW NATURE PROD HOLDING FRANCE SAS.
 XX Prevots F, Tolou S;
 XX WPI, 2001-611505/70.
 XX N-PSDB; AA165682.
 XX New nucleic acid for imparting phage resistance to lactic acid bacteria used in food preparation comprises a polynucleotide that encodes a phage-resistance mechanism -
 XX
 XX Disclosure; Page 19-20; 30pp; French.
 XX The present protein is encoded by a nucleic acid sequence which comprises a phage resistance mechanism. Plasmids that contain the phage resistance CC mechanism nucleic acids are used to impart phage resistance to CC industrially important strains of lactic acid bacteria. These recombinant CC strains are then used for preparation of cheese, yogurt, butter, sauces and sauerkraut.
 XX
 XX Sequence 260 AA;
 SQ
 Query Match 70.5%; Score 31; DB 22; Length 260;
 Best Local Similarity 62.5%; Pred. No. 2.5e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTOQYT 8
 Db 48 NEGLEQYT 55
 RESULT 22
 ID AAR39573
 AAR39573 standard; Protein; 667 AA.
 XX AAR39573;
 DT 25-MAR-2003 (updated)
 XX 07-FEB-1994 (first entry)
 DE Sequence of 741 SFV-PE40.
 XX Single chain Fv polypeptide; VH-VL heterodimer; immunoglobulin; Ig;
 KW exotoxin.
 KM Pseudomonas.
 XX OS
 XX WO9316185-A2.
 XX 19-AUG-1993.
 XX 05-FEB-1993; 93WO-US01055.
 XX 06-FEB-1992; 92US-0831967.
 XX (CETU) CETUS ONCOLOGY CORP.
 XX (CREA-) CREATIVE BIOMOLECULES INC.
 PI Houston LL, Huston JS, Oppermann H, Ring DB;
 DR WPI, 1993-272889/34.

DR N-PSDB; AAQ46088.
 XX New single chain Fv polypeptide binding to C-erbB-2 tumour
 PT antigen - for imaging or treating breast or ovarian cancer etc.
 XX
 XX Example; pages 71-74; 87pp; English.
 XX c-erbB-2 refers to a protein antigen expressed on the surface of
 CC tumour cells, such as breast and ovarian tumour cells, which is an
 CC approx. 200,000 mol. wt. acidic glycoprotein having an isoelectric
 CC pt. of about 5.3 (see AAQ46083, AAR39568). A single chain Fv (SFV)
 CC is a covalently linked VH-VL heterodimer which is expressed from
 CC a gene fusion including VH- and VL- encoding genes connected by
 CC a peptide-encoding linker. Such linker sequences are set forth in
 CC AA residues 116-135 in AAR39569, which includes part of the 16 AA
 CC linker sequences in AAR39572. Using AAQ46084 for the 520c9 monoclonal
 CC antibody, a single chain polypeptide can be produced having a
 CC binding affinity for a C-erbB-2 related antigen. In another design,
 CC the pseudomonal exotoxin fragment analogous to ricin A chain, PE40,
 CC is fused to the carboxy terminus of the anti-c-erbB-2 741 SFV.
 CC the resulting 741F8 SFV-PE40 is a single-chain Fv-toxin fusion
 CC protein, which was constructed with an 18 residue short PB leader
 CC which initially was left on the protein. 'X' in AAR39573 refers to
 CC posn. of stop codon in cDNA.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 667 AA;
 SQ
 Query Match 70.5%; Score 31; DB 14; Length 667;
 Best Local Similarity 62.5%; Pred. No. 7.1e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTOQYT 8
 Db 420 NRGTOQNT 427
 RESULT 23
 ID AAG82884
 AAG82884 standard; Protein; 727 AA.
 XX AAG82884;
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:2862.
 XX
 XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
 KM vaccination; endocarditis.
 KW
 XX Staphylococcus epidermidis.
 XX WO200134809-A2.
 XX 17-MAY-2001.
 XX 09-NOV-2000; 2000WO-US30782.
 XX 09-NOV-1999; 99US-0164258.
 XX (GLAX) GLAXO GROUP LTD.
 XX Kimmerly WJ;
 XX WPI; 2001-316495/33.
 XX N-PSDB; AAH53734.
 DR Nucleic acid encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -
 XX Claim 18; Page 747; 2188pp; English.
 PS AAH52304 to AAH53970 represent nucleic acids (1) encoding polypeptides
 XX

CC (II), given in AAG81454 to AAG81120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH51971 to
 CC AAH51090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH51091 to
 CC AAH51098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 727 AA:

Query Match 70.5%; Score 31; DB 22; Length 727;
 Best Local Similarity 71.4%; Pred. No. 7.9e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOOY 7
 Db 202 NEGTKOY 208

RESULT 24
 ABP39440
 ID ABP39440 standard; Protein; 738 AA.

XX ABP39440;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4285.

KW Staphylococcus epidermidis; open reading frame; ORF, bacterial infection;

KM antibacterial; gene therapy.

XX Staphylococcus epidermidis.

XX US6380370-B1.

XX 30-APR-2002.

PF 13-AUG-1998; 98US-0134001.

PR 14-AUG-1997; 97US-055779P.

PR 08-NOV-1997; 97US-064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucetie-Stamm LA, Bush D;

XX WPI; 2002-381255/41.

DR N-PSDB; ABN91985.

PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis

XX polypeptide, useful for diagnosing and treating bacterial infections -

XX Disclosure; SEQ ID 4285; 267bp; English.

CC ABN90538 to ABN93174 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP31524 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to

CC screen for compounds able to interfere with the S. epidermidis life
 CC cycle or inhibit S. epidermidis infection.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.

XX SQ Sequence 738 AA;

Query Match 70.5%; Score 31; DB 23; Length 728;
 Best Local Similarity 71.4%; Pred. No. 7.9e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOOY 7
 Db 213 NEGTKOY 219

RESULT 25
 ABG04999
 ID ABG04999 standard; Protein; 1037 AA.

XX ABG04999;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #4990.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS69186.

PT New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity -

XX Claim 20; SEQ ID NO 35358, 103bp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridization probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. ABG00010-ABG03077 represent novel human

XX diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1037 AA;

Query Match 70.5%; Score 31; DB 22; Length 1037;
 Best Local Similarity 62.5%; Pred. No. 1.2e+03;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYR 8
 DB 585 NKGQDEFT 592

RESULT 26
 ABB62913
 ID ABB62913 standard; Protein; 1837 AA.
 XX ABB62913;

26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 15531.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-655680/75.

XX N-PSDB; ABL07016.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 interactions -

PS Disclosure; SEQ ID NO 15531; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1837 AA;

Query Match 70.5%; Score 31; DB 22; Length 1837;
 Best Local Similarity 75.0%; Pred. No. 2.2e+03;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYR 8
 DB 1205 NKGWQYR 1212

RESULT 27

AAV51863
 ID AAV51863 standard; peptide; 13 AA.

XX AAV51863;

XX 15-JUN-2000 (first entry)

XX Pancreatic elastase immunogenic peptide #6.

XX Immunogenic; pancreatic elastase; diagnosis; detection; isoenzyme;
 KW pancreas; pancreatitis; cystic fibrosis.

XX Unidentified.

XX DE19923892-A1.

XX 09-MAR-2000.

XX 25-MAY-1999; 99DE-1023892.

XX 08-SEP-1998; 98DE-1040900.

XX (PRIV-) PRIVATES INST BIOSERV GMBH.

XX Heinrich H, Kleiner R, Meyer U, Wagner H;

XX WPI; 2000-258170/23.

XX Diagnosing disorders of pancreatic function such as pancreatitis, by
 PT determining content of all pancreatic elastase isoenzymes in serum or
 PT feces -

XX Claim 5; Column 6; 6pp; German.

XX This invention describes a novel diagnostic method for detecting a
 CC disorder of pancreatic function which comprises determining the total
 CC content of all pancreatic elastases (isoenzymes) in serum, secretions
 CC or excretions. The method can be used to diagnose or monitor pancreatic
 CC disease such as chronic or acute pancreatitis. Test kits for detecting
 CC elastase-1 can also be used for diagnosis and monitoring of cystic
 CC fibrosis. The method is more sensitive than prior art tests that
 CC determine only elastase-1. Antibodies raised against immunogenic
 CC fragments of the different elastase isoforms are very specific and do
 CC not react with other components of feces. AAV51858-Y51872 represent
 CC pancreatic elastase immunogenic peptides described in the method of the
 CC invention.

XX Sequence 13 AA;

Query Match 68.2%; Score 30; DB 21; Length 13;
 Best Local Similarity 71.4%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYR 7
 DB 6 NDGTQOYR 12

RESULT 28

AAW40575
 ID AAW40575 standard; protein; 98 AA.

XX AAW40575;

XX 21-JUL-1998 (first entry)

XX Human IgG2 CH1 domain protein fragment.

XX Immunoglobulin G; IgG molecule; human; Fc region; LFA-1 receptor;
 KW disorder; salvage receptor binding epitope; cell adherence interaction;

KW Lymphocyte; T cell inflammatory response.
 XX Homo sapiens.
 OS US5739277-A.
 PN 14-APR-1998.
 PD 14-APR-1995; 95US-0422101.
 PF 14-APR-1995; 95US-0422101.
 PR 14-APR-1995; 95US-0422101.
 XX (GETH) GENENTECH INC.
 PA Presta LG, Snedecor BR;
 PI WPI, 1998-250490/22.
 DR Polypeptide(s) that are not Fc fragments and have an increased
 XX half-life - are useful for the treatment of LFA-1 mediated disorders
 PT Disclosure; Fig 2; 38pp; English.
 CC AAW40574-77 are fragments of human immunoglobulin G, (IgG) CH1 domains.
 CC These fragments are used in a novel method to produce polypeptides which
 CC contain an epitope from the Fc region of an IgG molecule and a mutated
 CC salvage receptor binding epitope. They are useful for the treatment of
 CC LFA-1 mediated disorders. These are conditions caused by cell adherence
 CC interactions involving the LFA-1 receptor on lymphocytes, e.g. T cell
 CC inflammatory responses. The mutated salvage receptor sequence in the
 CC polypeptides means that they have increased in vivo circulatory
 CC half-lives when compared to normal Fc regions of IgG molecules.
 SQ Sequence 98 AA;
 QY 1 NKGTQOYT 8
 DB 75 NFGTQTYT 82
 QY 1 NKGTQOYT 8
 DB 75 NFGTQTYT 82
 RESULT 29
 AAW08742
 ID AAY08742 standard; protein; 98 AA.
 AAY08742;
 DT 10-AUG-1999 (first entry)
 XX Human IgG3 protein CH1 domain.
 DE Human IgG3 protein CH1 domain.
 XX IgG; immunoglobulin G; CH1 domain; human; anti-CD18; IgG1; IgG2; IgG3;
 KW IgG4; kappa-CL domain; lambda-CL domain; focal ischaemic stroke;
 KW cerebroprotective; cerebral artery obstruction; blood flow; infarct;
 KW CD18 extracellular domain; endothelium; CD11b/CD18 complex dissociation;
 KW antibody.
 OS Homo sapiens.
 OS US5914112-A.
 PN 22-JUN-1999.
 PD 22-JUN-1997; 97US-0788800.
 PF 22-JUN-1997; 97US-0788800.
 PR 23-JAN-1996; 96US-0093038.
 PR 22-JAN-1997; 97US-0788800.
 XX (GETH) GENENTECH INC.
 PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.

XX Bednar MM, Gross CE, Thomas GR;
 PI WPI, 1999-370483/31.
 DR Anti-CD18 antibodies in stroke
 XX Disclosure; Fig 4A-B; 25pp; English.
 CC This invention describes a method for improving the clinical outcome in
 CC focal ischemic stroke by administering novel anti-CD18 antibody which
 CC has cerebroprotective properties. The invention particularly describes a
 CC method of treating focal ischemic stroke caused by the obstruction of a
 CC main cerebral artery which comprises administering an anti-CD18 antibody
 CC to increase the blood flow or reduce the infarct size, where: (1) the
 CC antibody binds to an extracellular domain of CD18 and inhibits or reduces
 CC the ability of the cell expressing CD18 to bind to endothelium, (2) the
 CC antibody binds CD18 with an affinity of less than 4 nm, or (3) the
 CC antibody dissociates CD11b/CD18 complex. This sequence represents the
 CC human IgG2 CH1 domain which is used to illustrate the method of the
 CC invention.
 SQ Sequence 98 AA;
 QY 1 NKGTQOYT 8
 DB 75 NFGTQTYT 82
 QY 1 NKGTQOYT 8
 DB 75 NFGTQTYT 82
 RESULT 30
 AAW92422
 ID AAW92422 standard; peptide; 98 AA.
 AC AAW92422;
 XX 23-APR-1999 (first entry)
 DT Human IgG2 protein CH1 domain.
 DE Human IgG2 protein CH1 domain.
 XX Antibody; salvage receptor binding epitope; Fab; F(ab')2; immunoglobulin;
 KW CH region; CL region; kidney; Fc region; CH1 domain; CH2 domain; IgG2;
 KW renal clearance rate; circulatory half-life.
 OS Homo sapiens.
 OS US5869046-A.
 PN 09-FEB-1999.
 PD 14-APR-1995; 95US-0422092.
 PF 14-APR-1995; 95US-0422092.
 PR 14-APR-1995; 95US-0422092.
 XX (GETH) GENENTECH INC.
 PA Presta LG, Snedecor BR;
 PI WPI, 1999-152694/13.
 DR Production of antibody fragments with reduced renal clearance - by
 PT introducing salvage receptor binding epitope into CH1 or CL region
 XX Disclosure; Column 53-54; 38pp; English.
 CC This invention describes a method for preparing a variant Fab or F(ab')2
 CC polypeptide having increased half-life in vivo, where the polypeptide
 CC contains an Ig or Ig-like domain comprising a CH1 and/or CL region, is
 CC cleared from the kidneys and does not contain an IgG Fc region. The
 CC method involves altering the polypeptide within the CH1 or CL region to

Matches	6;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	1	NKGTQOY	8						
,									
Db	75	NFGTQY	82						

RESULT 32
ABG31880
ID ABG31880 standard; protein, 98 AA.

AC ABG31880;
XX
DT 05-NOV-2002 (first entry)

XX Human, IgG2 CH1 domain; immunoglobulin G; cerebral blood flow;
KW infarct size; focal ischemic stroke; main cerebral artery;

KW acute ischaemic stroke; thrombolytic therapy; thromboembolic stroke.

OS Homo sapiens.

PN US2002081294-A1.

PD 27-JUN-2002.

XX 20-DEC-2000; 2000US-0811384.
PF
XX

PR 17-FEB-1999; 99US-0251652.

XX

XX

XX

XX

PT ischaemic stroke using anti-CD18 antibody and tissue plasminogen

PT stroke -

PS Disclosure; Fig 4; 27pp; English.

CC The invention relates to a method of increasing cerebral blood flow and/or reducing infarct size in focal ischaemic stroke caused by obstruction of a main cerebral artery in a human, comprising co-administering tissue plasminogen activator and anti-CD18 antibody about 3-5 hours after the

CC ischaemic stroke and to provide an alternative to thrombolytic therapy

CC therapy has been unsuccessful or is contra-indicated. The present
CC sequence represents the human IgG2 CH1 domain used in the method of the
CC invention.
XX
SQ Sequence 98 AA;

```

Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 NKGTOQT 8
    | | | | |
Db 75 NFGTOQTY 82

```

RESULT 33

ID AAW41393 standard; Protein, 116 AA.
XX

```

AC AAW41393;
XX
XX 02-JUN-1998 (first entry)
XX
XX Human heavy chain IGG2CH1' constant region.
XX
XX Anti-CEA antibody; carcinoembryonic antigen; 806.077 Ab; cancer therapy;
XX cancer diagnosis; complementarity determining region; heavy chain.
XX
XX Homo sapiens.
XX
XX WO9742329-A1.
XX
XX 13-NOV-1997.
XX
XX 29-APR-1997; 97WO-GB01165.
XX
XX 14-FEB-1997; 97GB-0003103.
XX 04-MAY-1996; 96GB-0009405.
XX
XX (ZENNE ) ZENNECA LTD.
XX
XX Copley CG, Edge MD, Emery SC;
XX
XX WPI; 1997-558987/51.
XX
XX N-PSDB; AAV17290.
XX
XX Anti-carcinoembryonic antigen antibody 806.077 Ab - used for
XX diagnosis and therapy of cancer
XX
XX Example 8; Page 106; 208pp; English.
XX
XX This sequence is the human heavy chain IGG2CH1' constant region.
XX The antibody was used to create a humanised antibody of the invention.
XX The antibody is an anti-CEA (carcinoembryonic antigen) antibody
XX (806.077 Ab). Host cells or transgenic organisms transformed with DNA
XX encoding the antibody, are used to make the antibody or conjugate. The
XX conjugate is used in a medicament suitable for intravenous
XX administration. The conjugate can be used for cancer therapy, selectively
XX killing tumour cells. The antibody can be used for in vivo or in vitro
XX diagnosis of cancer.
XX
XX Sequence 116 AA;
XX
XX Query Match 68.2%; Score 30; DB 18; Length 116;
XX Best Local Similarity 75.0%; Pred. No. 1.7e+02;
XX Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 NKGTQOQYT 8
XX | | | | |
XX 75 NFGTQTYT 82
XX
XX Db
XX
XX RESULT 34
XX AAY25391
XX ID AAY25391 standard; Protein; 116 AA.
XX
XX AAY25391;
XX
XX 07-SEP-1999 (first entry)
XX
XX Chinaeric antibody PQXB1/2 constant heavy chain HuIGG2CH1' protein.
XX
XX Parquet; antibody; light chain; herbicide; resistant; crop plant;
XX weed control; tolerant; diquat; photosynthesis inhibitor; photosystem I;
XX free radical; lipid peroxidation; electron transport; photosystem II;
XX vacuole; cell surface; cytotoxic; sensitive; PQXB1/2; HuIGG2CH1';
XX constant region; heavy chain.
XX
XX Synthetic.
XX
XX WO9932630-A1.

```

```

PD 01-JUL-1999.
XX
XX 15-DEC-1998; 98WO-GB03760.
XX
XX 19-DEC-1997; 97GB-0026955.
XX
XX (ZENNE ) ZENNECA LTD.
XX
XX Holt DC, Jones PG;
XX
XX WPI; 1999-405173/34.
XX
XX N-PSDB; AAX78845; AAX78846.
XX
XX Herbicide binding proteins and related polynucleotides
XX
XX Disclosure; Page 38-39; 60pp; English.
XX
XX This invention describes a novel herbicide binding protein which can
XX confer herbicide resistance activity. Crop plants, such as soybean,
XX cotton, tobacco, sugarbeet, oilseed rape, canola, flax, sunflower,
XX potato, tomato, alfalfa, lettuce, maize, wheat, sorghum, rye, bananas,
XX barley, oat, turf grass, forage grass, sugar cane, pea, field bean,
XX rice, pine, poplar, apple, grape, citrus or nut plants, transformed
XX with a herbicide binding protein gene are resistant to the herbicide.
XX Hence, weeds can be selectively controlled in a field of the transformed
XX crops. The plants are substantially resistant or tolerant to herbicides,
XX such as paraquat or diquat, that inhibit photosynthesis by accepting
XX electrons from photosystem I thus generating free radicals which cause
XX lipid peroxidation or by blocking electron transport in photosystem II.
XX The herbicide binding protein advantageously sequester the herbicide,
XX e.g. at the cell surface or in the vacuoles of a treated plant.
XX Sequestration at the cell surface prevents the entry of the herbicide
XX into the cell so that the herbicide cannot reach its intracellular target
XX and exert any significant cytotoxic effect. The herbicide binding protein
XX inhibits the mobility of the herbicide from the application site to the
XX whole plant preventing the herbicide reaching particularly sensitive
XX organs. Additionally, tolerant plants can be produced against herbicides
XX that have more than one target site.
XX
XX Sequence 116 AA;
XX
XX Query Match 68.2%; Score 30; DB 20; Length 116;
XX Best Local Similarity 75.0%; Pred. No. 1.7e+02;
XX Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 NKGTQOQYT 8
XX | | | | |
XX 75 NFGTQTYT 82
XX
XX Db
XX
XX RESULT 35
XX AAR41707
XX ID AAR41707 standard; Protein; 235 AA.
XX
XX AAR41707;
XX
XX 25-MAR-2003 (updated)
XX 20-OCT-1993 (first entry)
XX
XX Murine 128.1 VH/human gamma-2 CH1 encoded by plasmid pAH4625.
XX
XX Polymerase chain reaction; primer; PCR; amplify; murine; pAH4625;
XX heavy; light; chain; variable; constant; region; anti-human; pAH4807;
XX transferrin; receptor; antibody; brain; capillary; pAH4808; gamma-2;
XX endothelial cell; conjugate; neuropharmacological; gamma-3; gamma-4;
XX diagnostic; agent; tumour; AIDS; stroke; epilepsy; monoclonal;
XX Parkinsons disease; Alzheimers disease; SP2/0 cell.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FH 1..19
XX FT Peptide
XX
XX "note = "Leader peptide"

```

FT Region 20..50
FT /label= FR1
FT Region 51..55
FT /label= CDRL
FT Region 56..68
FT /label= FR2
FT Region 69..85
FT /label= CDR2
FT Region 86..117
FT /label= FR3
FT Region 118..126
FT /label= CDR3
FT Region 127..137
FT /label= FR4
FT Region 138..235
FT /label= CH1
XX WO9310819-A1.
XX 10-JUN-1993.
XX 24-NOV-1992; 92MO-US10206.
XX 26-NOV-1991; 91US-0800458.
XX (ALKE-) ALKERMES INC.
XX Frیدن PM;
XX WPI; 1993-196742/24.
XX N-PSDB; AAQ43846.
XX Antibody conjugates specific for transferrin receptor - used
XX for diagnosis and treatment of cancer, AIDS and neurological
XX disorders
XX Disclousure; Fig 17G-H; 151pp; English.
XX
XX The sequences given in AAR41707-09 are encoded by the expression vector
XX PAH4625. This vector represents the cloning of the human gamma
XX isotype, gamma-2, with the variable region of the murine monoclonal
XX antibody 128.1. This plasmid encodes a chimeric monoclonal antibody
XX in which the heavy chain (VH) is derived from a murine source and the
XX sequences encoding CH1, CH2 and CH3 are derived from a human source.
XX This vector, in combination with the chimeric light chain vector,
XX pAG4611 (see also AAQ43845), was transfected into SP2/0 cells and clones
XX were isolated. 128.1 is an anti-human transferrin receptor antibody
XX which binds to the transferrin receptor on brain capillary endothelial
XX cells. This antibody may be used in a conjugate in which it is linked
XX to a neuropharmaceutical or diagnostic agent. The conjugate may be
XX used to treat or prevent neurological disorders eg. brain tumours,
XX AIDS, stroke, epilepsy, Parkinsons and Alzheimers disease. It may
XX also be used for diagnostic methods.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 235 AA;
SQ
Query Match 68.2%; Score 30; DB 14; Length 235;
Best Local Similarity 75.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 NKGTQOQY 8
DB 212 NFGTQY 219

DT 03-FEB-1991 (first entry)
XX Human pancreatic elastase-I.
DE Elastase-I; lipoprotein metabolism.
XX Homo sapiens.
XX EP244189-A.
XX 04-NOV-1987.
XX 27-APR-1987; 87EP-0303711.
XX 26-APR-1986; 86JP-0097259.
XX (SANTY) SANKYO CO LTD.
XX Takiguchi Y, Tani T, Ohmine T, Furukawa H, Kawashima I;
XX WPI; 1987-308412/44.
XX N-PSDB; AAN70043.
XX Recombinant human pancreatic elastase-I - produced from DNA from
XX human pancreas obtained using porcine elastase-I as probe.
XX Claim 2; Page 46; 55pp; English.
XX
XX The protein sequence encodes human pancreas elastase-I, and may
XX have H, Met or all or part of Thr-Gln-Asp-Leu-Pro-Glu-
XX Thr-Asn-Ala-Arg at its N-terminal. Elastase is useful in
XX clinical studies eg restoration of elasticity and expandability
XX of arterial walls, reduction of serum lipid abnormality and
XX improvement of serum lipoprotein metabolism.
XX See also AAN70044, AAN70045 and AAN71331-6.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 240 AA;
SQ
Query Match 68.2%; Score 30; DB 8; Length 240;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 NKGTQOQY 7
DB 65 NDGTBOY 71

RESULT 37
AAM06179 standard; Protein; 250 AA.
ID AAM06179;
AC AAM06179;
DT 17-FEB-1997 (first entry)
XX
XX Humanised A5B57 Fd fragment.
XX ribonuclease; human; bovine; pancreatic; anti-tumour therapy; ADEPT;
XX mustard-ribonucleotide; antibody directed enzyme prodrug therapy;
XX anti-neoplastic; prodrug; reverse polarity; ion pair interaction;
XX polymerase chain reaction; PCR; HP-RNase; Fd; F(ab')2.
XX
XX Synthetic.
XX
XX WO9620011-A1.
XX 04-JUL-1996.
XX 21-DEC-1995; 95WO-GB02991.
XX 16-AUG-1995; 95GB-0016810.
PR

PR 23-DEC-1994; 94GB-0026192.
 XX (ZENE) ZENECA LTD.
 PA
 XX Blakey DC, Boyle FT, Davies DH, Eggelte HI, Heaton DW,
 PI Henham JF, Hennequin LFA, Marsham PR, Rabin BR, Slater AM,
 PI Tarragona-Fiol A, Taylorson CJ;
 XX WPI; 1996-321650/32.
 DR N-PSDB; AAT42509.
 XX
 PT Two component system for anti-tumour therapy - comprising targeting
 PT moiety linked to mutated enzyme which can transform an
 PT anti:neoplastic produg
 XX
 PS Example 6; Page 121-122; 182pp; English.
 XX
 CC A two-component system for anti-tumour therapy comprises a targeting
 CC moiety linked to a mutated enzyme which can transform an anti-neoplastic
 CC produg. The system is based on antibody directed enzyme produg therapy
 CC (ADEPT) using a non-naturally occurring mutant form of a host enzyme,
 CC pref. human pancreatic ribonuclease (HP-RNase), (see AAT42478-83). The
 CC targeting moiety can be an antibody, in partic. murine monoclonal
 CC antibody ASB7 (which binds to human carcinoembryonic antigen). ASB7 is
 CC suitable for targeting colorectal carcinoma. Fragments, esp. F(ab')₂,
 CC of the antibody can be conjugated to HP-RNase. ASB7 Fd and L chain
 CC fragments were isolated by PCR using cDNA isolated from ASB7 hybridoma
 CC cells. The present sequence is that of a humanised ASB7 Fd fragment.
 CC
 SQ Sequence 250 AA;
 XX
 Query Match 68.2%; Score 30; DB 17; Length 250;
 Best Local Similarity 75.0%; Pred. No. 3.9e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NKGTQOYT 8
 DB 215 NKGTOYTYT 222
 XX
 RESULT 38
 AAP70047
 ID AAP70047 standard; protein; 252 AA.
 XX
 AC AAP70047;
 XX
 DT 25-MAR-2003 (updated)
 03-FEB-1991 (first entry)
 XX
 Human chromosomal pancreas elastase-I.
 XX
 KW Elastase-I; lipoprotein metabolism.
 XX
 OS Homo sapiens.
 XX
 PN EP244189-A.
 XX
 PD 04-NOV-1987.
 XX
 PF 27-APR-1987; 87EP-0303711.
 XX
 PR 26-APR-1986; 86JP-0097259.
 XX
 PA (SANY) SANKYO CO LTD.
 XX
 PI Takiguchi Y, Tani T, Ohmine T, Furukawa H, Kawashima I;
 XX
 DR WPI; 1987-308412/44.
 DR N-PSDB; AAN70044, AAN71331, AAN71332, AAN71333, AAN71334, AAN71335,
 DR AAN71336.
 XX
 PT Recombinant human pancreatic elastase-I - produced from DNA from
 PT human pancreas obtained using porcine elastase-I as probe.

XX
 PS Disclosure; Page 20-24; 55pp; English.
 XX
 CC Human pancreatic elastase I is encoded by 7 exons. The amino acid
 CC sequence exhibits sequential similarity with rat elastase I and
 CC porcine elastase I of about 90 per cent. Elastase is
 CC useful in clinical studies, eg restoration of elasticity and
 CC expandability of arterial walls, reduction of serum lipid
 CC abnormality and improvement of serum lipoprotein metabolism.
 CC See also AAN70043, N70045, AAN71331-6.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 SQ Sequence 252 AA;
 XX
 Query Match 68.2%; Score 30; DB 8; Length 252;
 Best Local Similarity 71.4%; Pred. No. 3.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTQOY 7
 DB 77 NDGTQOY 83
 XX
 RESULT 39
 AAW41394
 ID AAW41394 standard; Protein; 255 AA.
 XX
 AC AAW41394;
 XX
 DT 02-JUN-1998 (first entry)
 XX
 DE Chimeric anti-CEA antibody 806.077 HuIgG2 Fd chain.
 XX
 KW Anti-CEA antibody; carcinoembryonic antigen; 806.077 Ab; cancer therapy;
 KW cancer diagnosis; complementarity determining region; Fd chain.
 XX
 OS Chimeric - Homo sapiens.
 OS Chimeric - Mus sp.
 XX
 PN WO9742329-A1.
 XX
 PD 13-NOV-1997.
 XX
 PF 29-APR-1997; 97WO-GB01165.
 XX
 PR 14-FEB-1997; 97GB-0003103.
 PR 04-MAY-1996; 96GB-0009405.
 XX
 PA (ZENE) ZENECA LTD.
 XX
 PI Copley CG, Edge MD, Emery SC,
 XX
 DR WPI; 1997-558987/51.
 DR N-PSDB; AAN71291.
 XX
 PT Anti-carcinoembryonic antigen antibody 806.077 Ab - used for
 PT diagnosis and therapy of cancer
 XX
 PS Example 8; Page 102-103; 208pp; English.
 XX
 CC This sequence is the HuIgG2 Fd chain region of the antibody of
 CC the invention. The antibody is an anti-CEA (carcinoembryonic antigen)
 CC antibody (806.077 Ab). Host cells or transgenic organisms transformed
 CC with DNA encoding the antibody, are used to make the antibody or
 CC conjugate. The conjugate is used in a medicament suitable for intravenous
 CC administration. The conjugate can be used for cancer therapy, selectively
 CC killing tumour cells. The antibody can be used for in vivo or in vitro
 CC diagnosis of cancer.
 CC
 SQ Sequence 255 AA;
 XX
 Query Match 68.2%; Score 30; DB 18; Length 255;
 Best Local Similarity 75.0%; Pred. No. 4e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
 |||||
 Db 214 NKGTQTYT 221

RESULT 40
 AAW41399
 ID AAW41399 standard; Protein; 255 AA.
 XX
 AC AAW41399;

DT 02-JUN-1998 (first entry)
 XX
 DE Humanised antibody 806.077 variable heavy chain.

KM Anti-CEA antibody; carcinoembryonic antigen; 806.077 Ab; cancer therapy;
 KW cancer diagnosis; complementarity determining region; heavy chain.

Chimeric - Homo sapiens.
 Chimeric - Mus sp.

XX
 FN WO9742329-A1.

XX
 PD 13-NOV-1997.

XX
 PE 29-APR-1997; 97WO-GB01165.

XX
 PR 14-FEB-1997; 97GB-0003103.

XX
 PR 04-MAY-1996; 96GB-0009405.

XX
 PA (ZENNE) ZENBECA LTD.

XX
 PI Copley CG, Edge MD, Emery SC;

XX
 DR WPI; 1997-558987/51.

XX
 DR N-PSDB; AAV17299.

XX
 PT Anti-carcinoembryonic antigen antibody 806.077 Ab - used for
 diagnosis and therapy of cancer

XX
 PS Example 11; Page 129-130; 208pp; English.

CC This sequence is the heavy chain variable region of the antibody of
 CC the invention. The antibody is an anti-CEA (carcinoembryonic antigen)
 CC antibody (806.077 Ab). Host cells or transgenic organisms transformed
 CC with DNA encoding the antibody, are used to make the antibody or
 CC conjugate. The conjugate is used in a medicament suitable for intravenous
 CC administration. The conjugate can be used for cancer therapy, selectively
 CC killing tumour cells. The antibody can be used for in vivo or in vitro
 CC diagnosis of cancer.

XX
 SQ Sequence 255 AA;

Query Match 68.2%; Score 30; DB 18; Length 255;
 Best Local Similarity 75.0%; Pred. NO. 4e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
 |||||
 Db 214 NKGTQTYT 221

Search completed: August 20, 2003, 09:29:38
 Job time : 29.2165 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using SW model

Run on: August 20, 2003, 09:19:07 ; Search time 8.82474 Seconds
(without alignments)
36.357 Million cell updates/sec

Title: US-09-991-433-6
Perfect score: 44
Sequence: 1 NKG2OQYT 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	395	US-08-856-841-13	Sequence 13, Appl
2	44	100.0	398	US-08-856-841-21	Sequence 21, Appl
3	44	100.0	415	US-08-856-841-20	Sequence 20, Appl
4	44	100.0	543	US-08-856-841-22	Sequence 22, Appl
5	44	100.0	756	US-09-438-268-4	Sequence 4, Appl
6	31	70.5	240	US-08-278-091-11	Sequence 11, Appl
7	31	70.5	240	US-08-483-859-11	Sequence 11, Appl
8	31	70.5	240	US-08-472-173-11	Sequence 11, Appl
9	31	70.5	240	US-08-487-167-11	Sequence 11, Appl
10	31	70.5	240	US-08-482-816-11	Sequence 11, Appl
11	31	70.5	240	US-08-296-149-11	Sequence 11, Appl
12	31	70.5	240	US-08-801-499-11	Sequence 11, Appl
13	31	70.5	240	US-08-615-271-11	Sequence 11, Appl
14	31	70.5	240	US-09-074-660-11	Sequence 11, Appl
15	31	70.5	240	US-09-074-659-11	Sequence 11, Appl
16	31	70.5	240	US-09-106-468-11	Sequence 11, Appl
17	31	70.5	240	US-09-106-466A-11	Sequence 11, Appl
18	31	70.5	240	US-09-106-467-11	Sequence 11, Appl
19	31	70.5	622	US-08-356-786-16	Sequence 16, Appl
20	31	70.5	630	US-09-328-352-7722	Sequence 1722, Ap
21	31	70.5	738	US-09-134-001C-4285	Sequence 4285, Ap
22	31	70.5	1068	US-08-390-874C-11	Sequence 11, Appl
23	31	70.5	1068	US-09-265-772-11	Sequence 11, Appl
24	30	68.2	48	5219996-6	Patent No. 5219996
25	30	68.2	98	US-08-422-101-5	Sequence 5, Appl
26	30	68.2	98	US-08-422-091-5	Sequence 5, Appl
27	30	68.2	98	US-08-422-092-5	Sequence 5, Appl

28	30	68.2	98	2	US-08-788-800-2	Sequence 2, Appl
29	30	68.2	98	3	US-08-422-093-5	Sequence 5, Appl
30	30	68.2	98	3	US-08-422-112-5	Sequence 5, Appl
31	30	68.2	116	3	US-09-171-945-22	Sequence 22, Appl
32	30	68.2	235	3	US-08-444-644-28	Sequence 28, Appl
33	30	68.2	235	4	US-08-232-246A-28	Sequence 28, Appl
34	30	68.2	250	4	US-09-011-769A-25	Sequence 25, Appl
35	30	68.2	255	4	US-09-171-945-19	Sequence 19, Appl
36	30	68.2	255	3	US-09-171-945-57	Sequence 57, Appl
37	30	68.2	326	2	US-08-656-586-9	Sequence 9, Appl
38	30	68.2	450	2	US-08-788-800-12	Sequence 12, Appl
39	30	68.2	469	2	US-07-934-373C-23	Sequence 23, Appl
40	30	68.2	469	3	US-08-437-642B-23	Sequence 23, Appl
41	30	68.2	469	4	US-08-146-206C-23	Sequence 23, Appl
42	30	68.2	530	3	US-08-477-460B-4	Sequence 4, Appl
43	30	68.2	530	3	US-08-379-516-4	Sequence 4, Appl
44	30	68.2	530	3	US-09-329-916-4	Sequence 4, Appl
45	30	68.2	530	3	US-08-485-372A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-856-841-13
Sequence 13, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856, 841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917, 096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 395
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
222

Query Match 100.0%; Score 44; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
245 NKGTQOY 252

RESULT 3

US-08-856-841-20
Sequence 20, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 415
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIOSUM)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 20:
US-08-856-841-20

Query Match 100.0%; Score 44; DB 3; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
256 NKGTQOY 263

RESULT 4

US-08-856-841-22
Sequence 22, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 543
TYPE: AMINO ACID

TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHEICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match 100.0%; Score 44; DB 3; Length 543;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTOOYT 8
|||||
Db 404 NKGTOOYT 411

RESULT 5
US-09-438-268-4
Sequence 4, Application US/09438268
Patent No. 6491907
GENERAL INFORMATION:
APPLICANT: Rabinowitz, Joseph E.
APPLICANT: Samulski, Richard J.
APPLICANT: Xiao, Weidong
TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
TITLE OF INVENTION: THE SAME
FILE REFERENCE: 5470-186
CURRENT APPLICATION NUMBER: US/09/438,268
CURRENT FILING DATE: 1999-11-10
EARLIER APPLICATION NUMBER: 60/107,840
EARLIER FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 60/123,651
EARLIER FILING DATE: 1999-03-10
NUMBER OF SEQ ID NOS: 59
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 4
LENGTH: 756
TYPE: PRT
ORGANISM: VITUS
US-09-438-268-4

Query Match 100.0%; Score 44; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTOOYT 8
|||||

Db 597 NKGTOOYT 604

RESULT 6
US-08-278-091-11
Sequence 11, Application US/08278091
Patent No. 5506139
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hinf7 Protein with
TITLE OF INVENTION: Reduced Protease Activity
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/278,091
FILING DATE: 21-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-371
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-278-091-11

Query Match 70.5%; Score 31; DB 1; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTOOYT 7
|||||
Db 65 NKGTOOYT 71

RESULT 7
US-08-483-859-11
Sequence 11, Application US/08483859
Patent No. 5656436
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hinf7 Protein with
TITLE OF INVENTION: Reduced Protease Activity
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto

STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,859
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-495 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-483-859-11

Query Match 70.5%; Score 31; DB 1; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOY 7
Db 65 NNGTEQY 71

RESULT 8
US-08-472-173-11
Sequence 11, Application US/08472173
Patent No. 5665353
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hinf7 Protein with
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,173
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149

FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-493 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-472-173-11

Query Match 70.5%; Score 31; DB 1; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOY 7
Db 65 NNGTEQY 71

RESULT 9
US-08-487-167-11
Sequence 11, Application US/08487167
Patent No. 5869302
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hinf7 Protein with
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,167
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-508 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-487-167-11

Query Match 70.5%; Score 31; DB 2; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
| | | | |
DB 65 NNGTEOY 71

RESULT 10
US-08-482-816-11
Sequence 11, Application US/08482816
Patent No. 5935573

GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Act
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,816
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994

ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-494 MIS.VG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-482-816-11

Query Match 70.5%; Score 31; DB 2; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
| | | | |
DB 65 NNGTEOY 71

RESULT 11

US-08-296-149-11
Sequence 11, Application US/08296149
Patent No. 5939297

GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
Reduced Protease Activity
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,149
FILING DATE:

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-390
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-296-149-11

Query Match 70.5%; Score 31; DB 2; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
| | | | |
DB 65 NNGTEOY 71

RESULT 12
US-08-801-499-11
Sequence 11, Application US/08801499
Patent No. 5962430

GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Act
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,499
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/482,816
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-671 MTS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-801-499-11

Query Match 70.5%; Score 31; DB 2; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOY 7
Db 65 NNGTEQY 71

RESULT 13
US-08-615-271-11
Sequence 11, Application US/08615271
Patent No. 5981503
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michael H.
TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
TITLE OF INVENTION: PROTEASE ACTIVITY
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,271
FILING DATE: 20-JUN-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-580

TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-615-271-11

Query Match 70.5%; Score 31; DB 2; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOY 7
Db 65 NNGTEQY 71

RESULT 14
US-09-074-660-11
Sequence 11, Application US/09074660
Patent No. 6020183
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michael H.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
TITLE OF INVENTION: Reduced Protease Activity
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/074,660
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,167
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-731 MTS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-074-660-11

Query Match 70.5%; Score 31; DB 3; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
| | | | |
| | | | |
DB 65 NNGTEOY 71

RESULT 15

US-09-074-659-11
; Sequence 11, Application US/09074659
; Patent No. 6025342
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michael H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,659
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-730 MIS:jfb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1153
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-074-659-11

Query Match 70.5%; Score 31; DB 3; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
| | | | |
| | | | |
DB 65 NNGTEOY 71

RESULT 16
US-09-106-468-11

; Sequence 11, Application US/09106468
; Patent No. 6114125
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michael H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,468
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-825
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1153
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-106-468-11

Query Match 70.5%; Score 31; DB 3; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
| | | | |
| | | | |
DB 65 NNGTEOY 71

RESULT 17
US-09-106-466A-11
; Sequence 11, Application US/09106466A
; Patent No. 6147057
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michael H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario

US-09-106-466A-11
; Sequence 11, Application US/09106466A
; Patent No. 6147057
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michael H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario

COUNTRY: Canada
ZIP: MSG 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/106,466A
FILING DATE:
CLASSIFICATION: 514514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/615,271
FILING DATE: 20-JUN-1996
CLASSIFICATION: 514514
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-826
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-106-466A-11

Query Match 70.5% Score 31; DB 3; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOY 7
Db 65 NNGTEQY 71

RESULT 18
US-09-106-467-11
Sequence 11, Application US/09106467
Patent No. 6153580
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: KOENEN, Raymond P.
APPLICANT: KLEIN, Michael H.
TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
TITLE OF INVENTION: PROTEASE ACTIVITY
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: MSG 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/106,467
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/615,271
FILING DATE: 20-JUN-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-824
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-106-467-11

Query Match 70.5% Score 31; DB 3; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOY 7
Db 65 NNGTEQY 71

RESULT 19
US-08-356-786-16
Sequence 16, Application US/08356786
Patent No. 5877305
GENERAL INFORMATION:
APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Teesta, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 622 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-786-16

Query Match 70.5% Score 31; DB 2; Length 622;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
:|:|:|:|
DB 420 NRGTONMT 427

RESULT 20
US-09-328-352-7722
; Sequence 7722, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; PENDING FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7722
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7722

Query Match 70.5%; Score 31; DB 4; Length 630;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
:|:|:|:|
DB 186 NRGCCOYT 193

RESULT 21
US-09-134-001C-4285
; Sequence 4285, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PENDING FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; ID NO 4285
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4285

Query Match 70.5%; Score 31; DB 4; Length 738;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYT 7
:|:|:|:|
DB 213 NRGTRQY 219

RESULT 22
US-08-390-874C-11
; Sequence 11, Application US/08390874C
; Patent No. 6043062
; GENERAL INFORMATION:
; APPLICANT: Kippel, Anke
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: A Constitutively Active
; TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof
; NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,874C
FILING DATE: 17-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 02307K-057000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-874C-11

Query Match 70.5%; Score 31; DB 3; Length 1068;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
:|:|:|:|
DB 965 SKGQNEYT 972

RESULT 23
US-09-265-772-11
; Sequence 11, Application US/09265772
; Patent No. 6300111
; GENERAL INFORMATION:
; APPLICANT: Kippel, Anke
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: A Constitutively Active
; TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof
; NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,772
FILING DATE: 10-MAR-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,874
FILING DATE: 17-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684

REFERENCE/DOCKET NUMBER: 02307K-057010US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-265-772-11

Query Match
Best Local Similarity 70.5%; Score 31; DB 4; Length 1068;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
695 SKGAQBYT 972

RESULT 24
5219996-6
Patent No. 5219996
APPLICANT: BODMER, MARK W.; ADAIR, JOHN R.; WHITTLE, NIGEL R.
LYONS, ALAN H.; OWENS, RAYMOND J.
TITLE OF INVENTION: RECOMBINANT ANTIBODIES AND METHODS FOR
THEIR PRODUCTION IN WHICH SURFACE RESIDUES ARE ALTERED TO
CYSTEINE RESIDUES FOR ATTACHMENT OF EFFECTOR OR RECEPTOR
MOLECULES
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/353,634
FILING DATE: 05-SEP-1988
SEQ ID NO: 6
LENGTH: 48
5219996-6

Query Match
Best Local Similarity 68.2%; Score 30; DB 6; Length 48;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
Db 25 NFGTQY 32

PLT 25
08-422-101-5
Sequence 5, Application US/08422101
Patent No. 5739277
GENERAL INFORMATION:
APPLICANT: Leonard Presta
APPLICANT: Brad Snedecor
TITLE OF INVENTION: Altered Polypeptides with Increased
TITLE OF INVENTION: Half-Life
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 KB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,101
FILING DATE: 14-APR-1995

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 932-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-422-101-5

Query Match
Best Local Similarity 68.2%; Score 30; DB 1; Length 98;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
Db 75 NFGTQY 82

RESULT 26
US-08-422-091-5
Sequence 5, Application US/08422091
Patent No. 5747035
GENERAL INFORMATION:
APPLICANT: Leonard Presta
APPLICANT: Brad Snedecor
TITLE OF INVENTION: Altered Polypeptides with Increased
TITLE OF INVENTION: Half-Life
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 KB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,091
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 932-6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-422-091-5

Query Match
Best Local Similarity 68.2%; Score 30; DB 1; Length 98;

Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
| | | | |
Db 75 NFGTQTYT 82

RESULT 27

US-08-422-092-5
Sequence 5, Application US/08422092
Patent No. 5869046
GENERAL INFORMATION:
APPLICANT: Leonard Presta
APPLICANT: Brad Snedecor
TITLE OF INVENTION: Altered Polypeptides with Increased
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,092
FILING DATE: 14-APR-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 932-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-422-092-5

Query Match 68.2%; Score 30; DB 2; Length 98;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
| | | | |
Db 75 NFGTQTYT 82

RESULT 28

US-08-788-800-2
Sequence 2, Application US/08788800
Patent No. 5914112
GENERAL INFORMATION:
APPLICANT: Bednar, Martin M.
APPLICANT: Thomas, G. Roger
APPLICANT: Gross, Cordell B.
TITLE OF INVENTION: ANTI-CD18 ANTIBODIES IN STROKE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/788,800
FILING DATE: 22-Jan-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0987r1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
US-08-788-800-2

Query Match 68.2%; Score 30; DB 2; Length 98;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
| | | | |
Db 75 NFGTQTYT 82

RESULT 29

US-08-422-093-5
Sequence 5, Application US/08422093
Patent No. 6096871
GENERAL INFORMATION:
APPLICANT: Leonard Presta
APPLICANT: Brad Snedecor
TITLE OF INVENTION: Altered Polypeptides with Increased
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,093
FILING DATE: 14-APR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 932
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-422-093-5

Query Match 68.2%; Score 30; DB 3; Length 98;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
Db 75 NFGTQTYT 82

RESULT 30
US-08-422-112-5
Sequence 5, Application US/08422112
Patent No. 6121022

GENERAL INFORMATION:
APPLICANT: Leonard Presta
APPLICANT: Brad Snedecor
TITLE OF INVENTION: Altered Polypeptides with Increased
TITLE OF INVENTION: Half-Life
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patlin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,112
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 932-1.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-422-112-5

Query Match 68.2%; Score 30; DB 3; Length 98;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
Db 75 NFGTQTYT 82

RESULT 31
US-09-171-945-22

Sequence 22, Application US/09171945
Patent No. 6277599
GENERAL INFORMATION:
APPLICANT: Emery, Stephen
APPLICANT: Copley, Clive Graham
APPLICANT: Edge, Michael Derek

TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said
TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System
FILE REFERENCE: Monoclonal Antibody to CEA
CURRENT APPLICATION NUMBER: US/09/171,945
CURRENT FILING DATE: 1998-10-29

PRIOR APPLICATION NUMBER: GB9703103.3
PRIOR FILING DATE: 1997-02-14
PRIOR APPLICATION NUMBER: GB9609405.7
PRIOR FILING DATE: 1996-05-04
PRIOR APPLICATION NUMBER: PCT/GB97/01165
PRIOR FILING DATE: 1997-04-29
NUMBER OF SEQ ID NOS: 131
SOFTWARE: Patentlin Ver. 2.1
SEQ ID NO 22

LENGTH: 116
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: humanized

US-09-171-945-22

Query Match 68.2%; Score 30; DB 3; Length 116;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
Db 75 NFGTQTYT 82

RESULT 32
US-08-444-644-28

Sequence 28, Application US/08444644
Patent No. 6015555

GENERAL INFORMATION:
APPLICANT: Fiden, Phillip M.
TITLE OF INVENTION: TRANSFERRIN RECEPTOR SPECIFIC
TITLE OF INVENTION: ANTIBODY-NEUROPHARMACEUTICAL OR DIAGNOSTIC AGENT
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA

ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentlin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,644
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,246
FILING DATE: 07-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/800,458
FILING DATE: 26-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US90/05077
FILING DATE: 07-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/404,089

FILING DATE: 07-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Wagner, Richard W.
REGISTRATION NUMBER: 34,480
REFERENCE/DOCKET NUMBER: ALK88-15AAAZ
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
US-08-444-644-28

Query Match 68.2%; Score 30; DB 3; Length 235;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 NKGTQYTT 8
| | | | |
212 NFGTQYTT 219

RESULT 33
US-08-232-246A-28
Sequence 28, Application US/08232246A
Patent No. 6329508
GENERAL INFORMATION:
APPLICANT: Fiden, Phillip M.
TITLE OF INVENTION: TRANSFERRIN RECEPTOR SPECIFIC
TITLE OF INVENTION: ANTIBODY-NEUROPHARMACEUTICAL OR DIAGNOSTIC AGENT
TITLE OF INVENTION: CONJUGATES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitta Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,246A
FILING DATE: 04-MAY-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/800,458
FILING DATE: 26-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US90/05077
FILING DATE: 07-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/404,089
FILING DATE: 07-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Wagner, Richard W.
REGISTRATION NUMBER: 34,480
REFERENCE/DOCKET NUMBER: ALK88-15AAAZ
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
US-08-232-246A-28

Query Match 68.2%; Score 30; DB 4; Length 235;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQYTT 8
| | | | |
DB 212 NFGTQYTT 219

RESULT 34
US-09-011-769A-25
Sequence 25, Application US/09011769A
Patent No. 6436691
GENERAL INFORMATION:
APPLICANT: Slater, Anthony M.
BLAKEY, David C.
DAVIES, David H.
HENNAM, John F.
HENNEQUIN, Laurent F.A.
MARSHAM, Peter R.
DOWELL, Robert I.
TITLE OF INVENTION: Chemical Compounds
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSER: Pillsbury Madison & Sutro, LLP
STREET: 1100 New York Ave., N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 Mb disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/011,769A
FILING DATE: 13-Feb-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/01975
FILING DATE: 13-AUG-1996
APPLICATION NUMBER: GB 9612295.7
FILING DATE: 12-JUN-1996
APPLICATION NUMBER: GB 9611019.2
FILING DATE: 25-MAY-1996
APPLICATION NUMBER: GB 9516010.0
FILING DATE: 16-AUG-1995
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-011-769A-25

Query Match 68.2%; Score 30; DB 4; Length 250;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQYTT 8
| | | | |
DB 215 NFGTQYTT 222

RESULT 35
US-09-171-945-19
Sequence 19, Application US/09171945

Patent No. 6277599
GENERAL INFORMATION:
APPLICANT: Emery, Stephen
APPLICANT: Copley, Clive Graham
APPLICANT: Edge, Michael Derek
TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said
TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System
FILE REFERENCE: Monoclonal Antibody to CEA
CURRENT APPLICATION NUMBER: US/09/171,945
CURRENT FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: GB9703103.3
PRIOR FILING DATE: 1997-02-14
PRIOR APPLICATION NUMBER: GB9609405.7
PRIOR FILING DATE: 1996-05-04
PRIOR APPLICATION NUMBER: PCT/GB97/01165
PRIOR FILING DATE: 1997-04-29
NUMBER OF SEQ ID NOS: 131
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 19
LENGTH: 255
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: humanized
US-09-171-945-19

Query Match 68.2%; Score 30; DB 3; Length 255;
Best Local Similarity 75.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
Db 214 NFGTQYTY 221

RESULT 36
US-09-171-945-57
Sequence 57, Application US/09171945
Patent No. 6277599
GENERAL INFORMATION:
APPLICANT: Emery, Stephen
APPLICANT: Copley, Clive Graham
APPLICANT: Edge, Michael Derek
TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said
TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System
FILE REFERENCE: Monoclonal Antibody to CEA
CURRENT APPLICATION NUMBER: US/09/171,945
CURRENT FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: GB9703103.3
PRIOR FILING DATE: 1997-02-14
PRIOR APPLICATION NUMBER: GB9609405.7
PRIOR FILING DATE: 1996-05-04
PRIOR APPLICATION NUMBER: PCT/GB97/01165
PRIOR FILING DATE: 1997-04-29
NUMBER OF SEQ ID NOS: 131
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 57
LENGTH: 255
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: humanized
US-09-171-945-57

Query Match 68.2%; Score 30; DB 3; Length 255;
Best Local Similarity 75.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
Db 214 NFGTQYTY 221

RESULT 37
US-08-656-586-9
Sequence 9, Application US/08656586
Patent No. 5834597
GENERAL INFORMATION:
APPLICANT: Tso, J. Yun
APPLICANT: Cole, Michael S.
APPLICANT: Anasetti, Claudio
TITLE OF INVENTION: Mutated No. 5834597 inactivating IgG2 Domains and
TITLE OF INVENTION: Anti-CD3 Antibodies Incorporating the Same
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3634
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,586
FILING DATE: 31-MAY-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joseph O.
REGISTRATION/DOCKET NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 11823-007210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 326 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..326
OTHER INFORMATION: /note="heavy chain constant region of
OTHER INFORMATION: IgG2 mutant 3"
US-08-656-586-9

Query Match 68.2%; Score 30; DB 2; Length 326;
Best Local Similarity 75.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
Db 75 NFGTQYTY 82

RESULT 38
US-08-788-800-12
Sequence 12, Application US/08788800
Patent No. 5914112
GENERAL INFORMATION:
APPLICANT: Bednar, Martin M.
APPLICANT: Thomas, G. Roger
APPLICANT: Gross, Cordell E.
TITLE OF INVENTION: ANTI-CD18 ANTIBODIES IN STROKE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA

ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/788,800
FILING DATE: 22-Jan-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0387x1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-788-800-12

Query Match
Best Local Similarity 68.2%; Score 30; DB 2; Length 450;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYT 8
DB 199 NFGTQTYT 206

RESULT 39
US-07-934-373C-23
Sequence 23, Application US/07934373C
Patent No. 5821337
GENERAL INFORMATION:
APPLICANT: Paul J. Carter
TITLE OF INVENTION: Immunoglobulin Variants
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/934,373C
FILING DATE: 21-Aug-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05126
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/715272
FILING DATE: 14-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0709P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-07-934-373C-23

Query Match
Best Local Similarity 68.2%; Score 30; DB 2; Length 469;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYT 8
DB 218 NFGTQTYT 225

RESULT 40
US-08-437-642B-23
Sequence 23, Application US/08437642B
Patent No. 6054297
GENERAL INFORMATION:
APPLICANT: Paul J. Carter
TITLE OF INVENTION: Immunoglobulin Variants
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/437,642B
FILING DATE: 09-May-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/934373
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/146206
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05126
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/715272
FILING DATE: 14-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0709P2C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-437-642B-23

Query Match
Best Local Similarity 75.0%; Score 30; DB 3; Length 469;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYT 8
DB 218 NFGTQTYT 225

Wed Aug 20 12:54:16 2003

us-09-991-433-6.rai

Page 17

Search completed: August 20, 2003, 09:34:29
Job time : 9.82474 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:30:58 ; Search time 32.4948 Seconds
(without alignments)
32.488 Million cell updates/sec

Title: US-09-991-433-6
Perfect score: 44
Sequence: 1 NKGTQOYR 8

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 497079 seqs, 131961718 residues

1 number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PTC_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PTC_NEW_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	756	US-10-205-942-4	Sequence 4, Appl1
2	33	75.0	175	US-10-062-254-234	Sequence 234, App
3	31	70.5	239	US-09-910-071-15	Sequence 15, Appl
4	30	68.2	98	US-09-811-384-2	Sequence 22, Appl
5	30	68.2	116	US-09-910-059-22	Sequence 22, Appl
6	30	68.2	255	US-09-910-059-19	Sequence 19, Appl
7	30	68.2	255	US-09-910-059-57	Sequence 57, Appl
8	30	68.2	326	US-10-047-542-22	Sequence 22, Appl
9	30	68.2	443	US-09-256-156-2	Sequence 2, Appl1
10	30	68.2	450	US-09-811-384-12	Sequence 12, Appl
11	30	68.2	451	US-10-153-382-17	Sequence 17, Appl
12	30	68.2	461	US-10-249-011A-24	Sequence 24, Appl
13	30	68.2	463	US-10-153-382-3	Sequence 3, Appl1
14	30	68.2	463	US-10-153-382-5	Sequence 5, Appl1
15	30	68.2	463	US-10-153-382-13	Sequence 13, Appl

16	30	68.2	464	US-10-153-382-9	Sequence 9, Appl1
17	30	68.2	470	US-09-859-053-28	Sequence 28, Appl
18	30	68.2	470	US-09-859-053-32	Sequence 32, Appl
19	30	68.2	470	US-09-859-053-36	Sequence 36, Appl
20	30	68.2	472	US-10-006-593-67	Sequence 67, Appl
21	30	68.2	530	US-08-485-163-5	Sequence 5, Appl1
22	30	68.2	530	US-09-766-995-4	Sequence 4, Appl1
23	30	68.2	650	US-09-854-875A-1	Sequence 1, Appl1
24	30	68.2	650	US-10-056-253-1	Sequence 1, Appl1
25	29	65.9	32	US-09-864-761-15647	Sequence 45647, A
26	29	65.9	34	US-10-218-102-84	Sequence 84, Appl
27	29	65.9	248	US-10-002-784A-25	Sequence 25, Appl
28	29	65.9	368	US-10-284-400-4	Sequence 4, Appl1
29	29	65.9	397	US-10-284-400-12	Sequence 12, Appl
30	29	65.9	398	US-10-002-784A-24	Sequence 24, Appl
31	29	65.9	424	US-10-174-693-328	Sequence 328, App
32	29	65.9	468	US-10-002-784A-27	Sequence 27, Appl
33	29	65.9	514	US-10-021-425-28	Sequence 28, Appl
34	29	65.9	666	US-09-964-992A-5	Sequence 5, Appl1
35	29	65.9	711	US-09-964-992A-1	Sequence 1, Appl1
36	29	65.9	713	US-09-964-992A-2	Sequence 2, Appl1
37	29	65.9	717	US-09-854-122-49	Sequence 49, Appl
38	28	63.6	34	US-10-218-102-103	Sequence 103, App
39	28	63.6	34	US-10-218-102-226	Sequence 226, App
40	28	63.6	34	US-10-218-102-238	Sequence 238, App
41	28	63.6	36	US-09-864-761-35563	Sequence 35563, A
42	28	63.6	46	US-10-106-698-8209	Sequence 8209, Ap
43	28	63.6	87	US-09-864-761-48861	Sequence 48861, A
44	28	63.6	142	US-10-092-243A-19	Sequence 19, Appl
45	28	63.6	144	US-09-764-877-1818	Sequence 1818, Ap

ALIGNMENTS

```
RESULT 1
US-10-205-942-4
; Sequence 4, Application US/10205942
; Publication No. US2003005390A1
GENERAL INFORMATION:
; APPLICANT: University of No. US2003005390A1Ch Carolina-Chapel Hill
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulek, Richard J.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/10/205, 942
; CURRENT FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Adeno-associated virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2271)
; OTHER INFORMATION: B19/NAV chimeric capsid coding sequence
US-10-205-942-4
Query Match
Best Local Similarity 100.0%; Pred. No. 0.68; Length 756;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKGTQOYR 8
Db 597 NKGTQOYR 604
RESULT 2
US-10-062-254-234
; Sequence 234, Application US/10062254
; Publication No. US20020138882A1
```

GENERAL INFORMATION:
APPLICANT: Cahoon, Edgar B
APPLICANT: Cahoon, Rebecca B
APPLICANT: Falco, Saverio Carl
APPLICANT: Fang, Yiwen
APPLICANT: Hanke, Sabine S.
APPLICANT: Lee, Jian-Ming
APPLICANT: Li, Zhongsen
APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele
APPLICANT: Niu, Xiping
APPLICANT: Odeh, Joan
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
APPLICANT: Zheng, Peizhong
APPLICANT: Zhu, Qun
TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved in Plant Metabolism
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/062,254
CURRENT FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 09/630,346
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/146511
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/156006
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/156899
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 60/157287
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: 60/169767
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/171054
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/172958
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/171515
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 60/173535
PRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 375
SOFTWARE: Microsoft Office 97
SEQ ID NO: 234
LENGTH: 175
TYPE: PRT
ORGANISM: Trifolium aestivum
US-10-062-254-234
Query Match 75.0%; Score 33; DB 14; Length 175;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NKGTOOYT 8
DB 77 NKGTRLYT 84
RESULT 3
US-09-910-071-15
Sequence 15, Application US/09910071
Patent No. US20020116146A1
GENERAL INFORMATION:
APPLICANT: Tomikawa, Mayumi
APPLICANT: Atkawa, Seiji
APPLICANT: Matsuzawa, Fumiko
TITLE OF INVENTION: Method and Apparatus for Extracting and Evaluating Mutually Simi
TITLE OF INVENTION: Portions in One-Dimensional Sequences in Molecules and/or Three-
FILE REFERENCE: 522,1921D2
CURRENT APPLICATION NUMBER: US/09/910,071
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 08/014,867
PRIOR FILING DATE: 1993-02-08

NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 15
LENGTH: 239
TYPE: PRT
ORGANISM: unknown
FEATURE:
OTHER INFORMATION: amino acid sequence of Elastase
US-09-910-071-15
Query Match 70.5%; Score 31; DB 10; Length 239;
Best Local Similarity 71.4%; Pred. No. 97;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NKGTOOYT 7
DB 65 NKGTRQY 71
RESULT 4
US-09-811-384-2
Sequence 2, Application US/09811384
Patent No. US20020081294A1
GENERAL INFORMATION:
APPLICANT: Bednar, Martin M.
Gross, Cordell E.
Thomas, G. Roger
TITLE OF INVENTION: ANTI-CD18 ANTIBODIES IN STROKE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/811,384
FILING DATE: 20-Dec-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/251652
FILING DATE: 17-FEB-2000
APPLICATION NUMBER: 08/788800
FILING DATE: 22-JAN-1997
APPLICATION NUMBER: 60/093038
FILING DATE: 23-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P1729C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-5530
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-811-384-2
Query Match 68.2%; Score 30; DB 9; Length 98;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 NKGTOOYT 8
DB 77 NKGTRLYT 84

Db 75 NFGTQTYT 82

RESULT 5

US-09-910-059-22
Sequence 22, Application US/09910059
Patent No. US20020142359A1
GENERAL INFORMATION:
APPLICANT: Copley, Clive G
APPLICANT: Edge, Michael Derek
TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said Antibody,
FILE REFERENCE: 1991-209
CURRENT APPLICATION NUMBER: US/09/910,059
PRIOR FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 09/171,945
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: PCT/GB97/01165
PRIOR FILING DATE: 1997-04-29
PRIOR APPLICATION NUMBER: GB 9703103.3
PRIOR FILING DATE: 1997-02-14
PRIOR APPLICATION NUMBER: GB9609405.7
PRIOR FILING DATE: 1996-05-04
NUMBER OF SEQ ID NOS: 131
SOFTWARE: PatentIn version 3.1
SEQ ID NO 22
LENGTH: 116
TYPE: PRT
ORGANISM: Homo sapiens
US-09-910-059-22

Query Match 68.2%; Score 30; DB 10; Length 116;
Best Local Similarity 75.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Cy 1 NKGTQOY 8
Db 75 NFGTQTYT 82

RESULT 6
US-09-910-059-19
Sequence 19, Application US/09910059
Patent No. US20020142359A1
GENERAL INFORMATION:
APPLICANT: Copley, Clive G
APPLICANT: Edge, Michael Derek
TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said Antibody,
FILE REFERENCE: 1991-209
CURRENT APPLICATION NUMBER: US/09/910,059
PRIOR FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 09/171,945
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: PCT/GB97/01165
PRIOR FILING DATE: 1997-04-29
PRIOR APPLICATION NUMBER: GB 9703103.3
PRIOR FILING DATE: 1997-02-14
PRIOR APPLICATION NUMBER: GB9609405.7
PRIOR FILING DATE: 1996-05-04
NUMBER OF SEQ ID NOS: 131
SOFTWARE: PatentIn version 3.1
SEQ ID NO 19
LENGTH: 255
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chimaeric Hulg2 Fd construct
US-09-910-059-19

Query Match 68.2%; Score 30; DB 10; Length 255;

Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;Cy 1 NKGTQOY 8
Db 214 NFGTQTYT 221

RESULT 7

US-09-910-059-57
Sequence 57, Application US/09910059
Patent No. US20020142359A1
GENERAL INFORMATION:
APPLICANT: Copley, Clive G
APPLICANT: Edge, Michael Derek
TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said Antibody,
FILE REFERENCE: 1991-209
CURRENT APPLICATION NUMBER: US/09/910,059
PRIOR FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 09/171,945
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: PCT/GB97/01165
PRIOR FILING DATE: 1997-04-29
PRIOR APPLICATION NUMBER: GB 9703103.3
PRIOR FILING DATE: 1997-02-14
PRIOR APPLICATION NUMBER: GB9609405.7
PRIOR FILING DATE: 1996-05-04
NUMBER OF SEQ ID NOS: 131
SOFTWARE: PatentIn version 3.1
SEQ ID NO 57
LENGTH: 255
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: complete humanised Fd heavy chain sequence
US-09-910-059-57

Query Match 68.2%; Score 30; DB 10; Length 255;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 NKGTQOY 8
Db 214 NFGTQTYT 221

RESULT 8

US-10-047-542-22
Sequence 22, Application US/10047542
Publication No. US20020168367A1
GENERAL INFORMATION:
APPLICANT: LARRICK, JAMES W.
APPLICANT: WYCOFF, KEITH L.
TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL
DISEASES
FILE REFERENCE: 030905.0004.C1P1
CURRENT APPLICATION NUMBER: US/10/047,542
PRIOR FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: PCT/US01/13932
PRIOR FILING DATE: 2001-04-28
PRIOR APPLICATION NUMBER: 60/200,298
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 22
LENGTH: 326
TYPE: PRT
ORGANISM: Homo sapiens
US-10-047-542-22

Query Match 68.2%; Score 30; DB 14; Length 326;

Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
Db 75 NFGTQYTY 82

RESULT 9

US-09-256-156-2
Sequence 2, Application US/09256156A
Publication No. US20030105294A1
GENERAL INFORMATION:
APPLICANT: GILLIES, Stephen D
APPLICANT: LO, Kin-Ming
APPLICANT: LAN, Yan
APPLICANT: MESOLOWSKI, John
TITLE OF INVENTION: Enhancing the Circulating Half-life of Antibody-based
FILE REFERENCE: LEX-003
CURRENT APPLICATION NUMBER: US/09/256,156A
CURRENT FILING DATE: 1999-02-24
EARLIER APPLICATION NUMBER: US 60/075,887
EARLIER FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 443
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: IGG-2 CHAIN C REGION
NAME/KEY: VARIANT
LOCATION: (1)..(117)
OTHER INFORMATION: The Xaa at positions 1 to 117 are non-conserved
US-09-256-156-2

Query Match 68.2%; Score 30; DB 11; Length 443;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
Db 192 NFGTQYTY 199

RESULT 10

US-09-811-384-12
Sequence 12, Application US/09811384
Patent No. US20020081294A1
GENERAL INFORMATION:
APPLICANT: Bednar, Martin M.
Thomas, G. Roger
Gross, Cordell E.
TITLE OF INVENTION: ANTI-CD18 ANTIBODIES IN STROKE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/811,384
FILING DATE: 20-Dec-2000

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/251652
FILING DATE: 17-FEB-2000
APPLICATION NUMBER: 08/788800
FILING DATE: 22-JAN-1997
APPLICATION NUMBER: 60/093038
FILING DATE: 23-JAN-1996
ATTORNEY/AGENT INFORMATION:

NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P1729C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-5530
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-811-384-12

Query Match 68.2%; Score 30; DB 9; Length 450;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
Db 199 NFGTQYTY 206

RESULT 11

US-10-153-382-17
Sequence 17, Application US/10153382
Publication No. US20030086930A1
GENERAL INFORMATION:
APPLICANT: PRIZER PRODUCTS INC.
TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
FILE REFERENCE: PC3019A
CURRENT APPLICATION NUMBER: US/10/153,382
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: 60/293042
PRIOR FILING DATE: 2001-05-23
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 17
LENGTH: 451
TYPE: PRT
ORGANISM: Homo sapiens
US-10-153-382-17

Query Match 68.2%; Score 30; DB 15; Length 451;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
Db 200 NFGTQYTY 207

RESULT 12

US-09-249-011A-24
Sequence 24, Application US/09249011A
Patent No. US20020176855A1
GENERAL INFORMATION:
APPLICANT: CO, MAN SUNG
VASQUEZ, MAXIMILIANO
APPLICANT: CARBENO, BEATRIZ
APPLICANT: CELINKER, ABBIE CHERYL
APPLICANT: COLLINS, MARY
APPLICANT: GOLDMAN, SAMUEL
APPLICANT: GRAY, GARY S.

APPLICANT: KNIGHT, ANDREA
APPLICANT: O'HARA, DENISE
APPLICANT: RUP, BONITA
APPLICANT: VEIDMAN, GERTRUDA M.
TITLE OF INVENTION: HUMANIZED IMMUNOGLOBULIN REACTIVE WITH B7-2 AND METHODS
FILE REFERENCE: 08702.0081-00000
CURRENT APPLICATION NUMBER: US/09/249,011A
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 461
TYPE: PRT
ORGANISM: Mus sp.
US-09-249-011A-24

Query Match 68.2%; Score 30; DB 10; Length 461;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 NKGTOOYT 8
| | | | |
210 NFGTQTYT 217

RESULT 13
US-10-153-382-3
Sequence 3, Application US/10153382
Publication No. US20030086930A1
GENERAL INFORMATION:
APPLICANT: PFIZER PRODUCTS INC.
TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
FILE REFERENCE: PC23019A
CURRENT APPLICATION NUMBER: US/10/153,382
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: 60/293042
PRIOR FILING DATE: 2001-05-23
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 463
TYPE: PRT
ORGANISM: Homo sapiens
US-10-153-382-3

Query Match 68.2%; Score 30; DB 15; Length 463;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 NKGTOOYT 8
| | | | |
212 NFGTQTYT 219

RESULT 14
US-10-153-382-5
Sequence 5, Application US/10153382
Publication No. US20030086930A1
GENERAL INFORMATION:
APPLICANT: PFIZER PRODUCTS INC.
TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
FILE REFERENCE: PC23019A
CURRENT APPLICATION NUMBER: US/10/153,382
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: 60/293042
PRIOR FILING DATE: 2001-05-23
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 463
TYPE: PRT
ORGANISM: Homo sapiens

US-10-153-382-5

Query Match 68.2%; Score 30; DB 15; Length 463;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 NKGTOOYT 8
| | | | |
212 NFGTQTYT 219

RESULT 15
US-10-153-382-13
Sequence 13, Application US/10153382
Publication No. US20030086930A1
GENERAL INFORMATION:
APPLICANT: PFIZER PRODUCTS INC.
TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
FILE REFERENCE: PC23019A
CURRENT APPLICATION NUMBER: US/10/153,382
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: 60/293042
PRIOR FILING DATE: 2001-05-23
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 463
TYPE: PRT
ORGANISM: Homo sapiens
US-10-153-382-13

Query Match 68.2%; Score 30; DB 15; Length 463;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 NKGTOOYT 8
| | | | |
212 NFGTQTYT 219

RESULT 16
US-10-153-382-9
Sequence 9, Application US/10153382
Publication No. US20030086930A1
GENERAL INFORMATION:
APPLICANT: PFIZER PRODUCTS INC.
TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
FILE REFERENCE: PC23019A
CURRENT APPLICATION NUMBER: US/10/153,382
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: 60/293042
PRIOR FILING DATE: 2001-05-23
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 464
TYPE: PRT
ORGANISM: Homo sapiens
US-10-153-382-9

Query Match 68.2%; Score 30; DB 15; Length 464;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 NKGTOOYT 8
| | | | |
213 NFGTQTYT 220

RESULT 17
US-09-859-053-28
Sequence 28, Application US/09859053
Patent No. US20020102658A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Tezuka, Katsunari
/ APPLICANT: Tezuka, Katsunari
/ APPLICANT: Hori, No. US20020102658A1uaki
/ TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A
/ TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE ALLIM AND
/ FILE REFERENCE: 06501-079001
/ CURRENT FILING DATE: 2001-05-16
/ CURRENT APPLICATION NUMBER: US/09/859,053
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: JP 2001-99508
/ PRIOR FILING DATE: 2000-05-18
/ PRIOR APPLICATION NUMBER: JP 2000-147116
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 28
/ LENGTH: 470
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-053-28
```

```
Query Match      68.2%; Score 30; DB 10; Length 470;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 NKGTOOYT 8
      | | | | |
Db      219 NFGTQTYT 226
```

```
RESULT 18
US-09-859-053-32
/ Sequence 33, Application US/09859053
/ Patent No. US20020102658A1
/ GENERAL INFORMATION:
/ APPLICANT: Tezuka, Katsunari
/ APPLICANT: Tezuka, Katsunari
/ APPLICANT: Hori, No. US20020102658A1uaki
/ TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A
/ TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE ALLIM AND
/ FILE REFERENCE: 06501-079001
/ CURRENT APPLICATION NUMBER: US/09/859,053
/ CURRENT FILING DATE: 2001-05-16
/ PRIOR APPLICATION NUMBER: JP 2001-99508
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: JP 2000-147116
/ PRIOR FILING DATE: 2000-05-18
/ SOFTWARE: FastSeq for Windows Version 4.0
/ NUMBER OF SEQ ID NOS: 43
/ SEQ ID NO 32
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-859-053-32
```

```
Query Match      68.2%; Score 30; DB 10; Length 470;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 NKGTOOYT 8
      | | | | |
Db      219 NFGTQTYT 226
```

```
RESULT 19
US-09-859-053-36
/ Sequence 36, Application US/09859053
/ Patent No. US20020102658A1
/ GENERAL INFORMATION:
/ APPLICANT: Tezuka, Katsunari
/ APPLICANT: Tezuka, Katsunari
```

```
/ APPLICANT: Hori, No. US20020102658A1uaki
/ TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A
/ TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE ALLIM AND
/ FILE REFERENCE: 06501-079001
/ CURRENT FILING DATE: 2001-05-16
/ CURRENT APPLICATION NUMBER: US/09/859,053
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: JP 2001-99508
/ PRIOR FILING DATE: 2000-05-18
/ PRIOR APPLICATION NUMBER: JP 2000-147116
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 36
/ LENGTH: 470
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-859-053-36
```

```
Query Match      68.2%; Score 30; DB 10; Length 470;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 NKGTOOYT 8
      | | | | |
Db      219 NFGTQTYT 226
```

```
RESULT 20
US-10-006-593-67
/ Sequence 67, Application US/10006593
/ Publication No. US20030049683A1
/ GENERAL INFORMATION:
/ APPLICANT: Bowdish, Katherine S.
/ APPLICANT: Frederickson, Shana
/ APPLICANT: Renshaw, Mark
/ TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
/ FILE REFERENCE: 1087-2
/ CURRENT APPLICATION NUMBER: US/10/006,593
/ CURRENT FILING DATE: 2001-12-05
/ PRIOR APPLICATION NUMBER: US 60/251,448
/ PRIOR FILING DATE: 2000-12-05
/ PRIOR APPLICATION NUMBER: US 60/288,889
/ PRIOR FILING DATE: 2001-05-04
/ PRIOR APPLICATION NUMBER: US 60/294,068
/ PRIOR FILING DATE: 2001-05-29
/ NUMBER OF SEQ ID NOS: 118
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 67
/ LENGTH: 472
/ TYPE: PRT
/ ORGANISM: artificial sequence
/ OTHER INFORMATION: Humanized antibody heavy chain
/ US-10-006-593-67
```

```
Query Match      68.2%; Score 30; DB 15; Length 472;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 NKGTOOYT 8
      | | | | |
Db      221 NFGTQTYT 228
```

```
RESULT 21
US-08-485-163-5
/ Sequence 5, Application US/08485163
/ Publication No. US20020098191A1
/ GENERAL INFORMATION:
/ APPLICANT: Beaudry, Gary A.
/ APPLICANT: Maddon, Paul J.
/ TITLE OF INVENTION: CD4-GAMMA2 CD4-IgG2 CHIMERAS
```


NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,163
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-1-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-485-163-5

Query Match
Best Local Similarity 68.2%; Score 30; DB 8; Length 530;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOQYT 8
Db 279 NFGTQTYT 286

RESULT 22
US-09-766-995-4
Sequence 4, Application US/09766995
Patent No. US20020052481A1
GENERAL INFORMATION:
APPLICANT: Graham P. Allaway et al.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONT
FILE REFERENCE: 2048/41215-CB/JPW/SHS
CURRENT APPLICATION NUMBER: US/09/766,995
CURRENT FILING DATE: 2001-01-22
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent in version 3.0
SEQ ID NO 4
LENGTH: 530
TYPE: PRT
ORGANISM: homo sapians
US-09-766-995-4

RESULT 23
US-09-854-875A-1
Sequence 1, Application US/09854875A
Patent No. US20020076796A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: White, David
APPLICANT: Macbeth, Kyle J.
TITLE OF INVENTION: 2786, A NOVEL HUMAN AMINOPEPTIDASE
FILE REFERENCE: 35800/208224
CURRENT APPLICATION NUMBER: US/09/854,875A
CURRENT FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: PCT/US00/31873
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 09/443,795
PRIOR FILING DATE: 1999-11-19
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 650
TYPE: PRT
ORGANISM: Homo sapiens
US-09-854-875A-1

Query Match
Best Local Similarity 68.2%; Score 30; DB 9; Length 650;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOQYT 8
Db 597 NQKQKQYT 604

RESULT 24
US-10-056-253-1
Sequence 1, Application US/10056253
Publication No. US20020127694A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: White, David
APPLICANT: Macbeth, Kyle J.
TITLE OF INVENTION: 2786, A NOVEL HUMAN AMINOPEPTIDASE
FILE REFERENCE: 5800-62
CURRENT APPLICATION NUMBER: US/10/056,253
CURRENT FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: US/09/443,795
PRIOR FILING DATE: 1999-11-19
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 650
TYPE: PRT
ORGANISM: Homo sapiens
US-10-056-253-1

Query Match
Best Local Similarity 68.2%; Score 30; DB 14; Length 650;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOQYT 8
Db 597 NQKQKQYT 604

RESULT 25
US-09-864-761-45647
Sequence 45647, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.

```

APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 45647
LENGTH: 32
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011255.3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.82
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.83
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.71
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.72
US-09-864-761-45647

Query Match          65.9%; Score 29; DB 9; Length 32;
Best Local Similarity 62.5%; Pred. No. 30;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 NKGTOOYT 8
      |||:|
Db      22 SKGGRQYT 29

RESULT 26
US-10-218-102-84
; Sequence 84, Application US/10218102
; Publication No. US20030130827A1
; GENERAL INFORMATION:
; APPLICANT: Bentzien, Joerg
; APPLICANT: Dahiyat, Bassil I.

```

```

APPLICANT: Desjarlais, John R.
APPLICANT: Hayes, Robert J.
APPLICANT: Vielmeter, Jost
TITLE OF INVENTION: Protein Design Automation for Protein Libraries
FILE REFERENCE: A-67229-11/RPT/RMS/RMK
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: US/10/218,102
PRIOR FILING DATE: 2002-08-10
PRIOR APPLICATION NUMBER: US 09/927,790
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: US 60/311,545
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: US 60/324,899
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/351,937
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 60/352,103
PRIOR FILING DATE: 2002-01-25
NUMBER OF SEQ ID NOS: 432
SOFTWARE: PatentIn version 3.1
SEQ ID NO 84
LENGTH: 34
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-10-218-102-84

Query Match          65.9%; Score 29; DB 16; Length 34;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 NKGTOOYT 8
      |||:|
Db      21 NKDTQDMT 28

RESULT 27
US-10-002-784A-25
; Sequence 25, Application US/10002784A
; Publication No. US20030036644A1
; GENERAL INFORMATION:
; /33
; APPLICANT: Ulrich, Robert G.
; TITLE OF INVENTION: Bacterial Superantigen Vaccines
; FILE REFERENCE: 003/233/SAP
; CURRENT APPLICATION NUMBER: US/10/002,784A
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776
; PRIOR FILING DATE: 97-06-25; 98-09-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 25
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: mutant streptococcal pyrogenic exotoxin-B
US-10-002-784A-25

Query Match          65.9%; Score 29; DB 15; Length 248;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 NKGTOOYT 8
      |||:|
Db      63 NKGLMDYT 70

RESULT 28
US-10-284-400-4
; Sequence 4, Application US/10284400
; Publication No. US20030143345A1
; GENERAL INFORMATION:

```

APPLICANT: Reddish, Mark A.
APPLICANT: Hu, Mary C.
APPLICANT: Walls, Michael A.
APPLICANT: Dale, James B.
TITLE OF INVENTION: MULTIVALENT STREPTOCOCCAL VACCINE COMPOSITIONS AND
TITLE OF INVENTION: METHODS FOR USE
FILE REFERENCE: 48112.413
CURRENT APPLICATION NUMBER: US/10/284,400
CURRENT FILING DATE: 2002-10-28
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 368
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Hybrid of Group A Streptococci
US-10-284-400-4

Query Match 65.9%; Score 29; DB 12; Length 368;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQ 6
DB 233 NKGTEQ 238

RESULT 29
US-10-284-400-12
Sequence 12, Application US/10284400
Publication No. US20030143245A1
GENERAL INFORMATION:
APPLICANT: Reddish, Mark A.
APPLICANT: Hu, Mary C.
APPLICANT: Walls, Michael A.
APPLICANT: Dale, James B.
TITLE OF INVENTION: MULTIVALENT STREPTOCOCCAL VACCINE COMPOSITIONS AND
TITLE OF INVENTION: METHODS FOR USE
FILE REFERENCE: 48112.413
CURRENT APPLICATION NUMBER: US/10/284,400
CURRENT FILING DATE: 2002-10-28
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 397
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Hybrid of Group A Streptococci
US-10-284-400-12

Query Match 65.9%; Score 29; DB 12; Length 397;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQ 6
DB 248 NKGTEQ 253

RESULT 30
US-10-002-784A-24
Sequence 24, Application US/10002784A
Publication No. US20030036644A1
GENERAL INFORMATION:
APPLICANT: Ulrich, Robert G.
APPLICANT: Ulrich, Robert G.
TITLE OF INVENTION: Bacterial Superantigen Vaccines
FILE REFERENCE: 003/233/SAP
CURRENT APPLICATION NUMBER: US/10/002,784A

CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776
PRIOR FILING DATE: 97-06-25; 98-09-01
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 24
LENGTH: 398
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: mutant streptococcal pyrogenic exotoxin B prosegment
US-10-002-784A-24

Query Match 65.9%; Score 29; DB 15; Length 398;
Best Local Similarity 62.5%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYT 8
DB 208 NKGKDYT 215

RESULT 31
US-10-174-693-328
Sequence 328, Application US/10174693
Publication No. US20030131373A1
GENERAL INFORMATION:
APPLICANT: Blockberg, Leonard N.
APPLICANT: Havukala, Ilkka
TITLE OF INVENTION: Materials and Methods for the
TITLE OF INVENTION: Modification of Plant Lignin Content
FILE REFERENCE: 11000.1003C5
CURRENT APPLICATION NUMBER: US/10/174,693
CURRENT FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: US 08/975,316
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: US 08/713,000
PRIOR FILING DATE: 1996-09-11
PRIOR APPLICATION NUMBER: US 09/169,789
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: US 09/615,192
PRIOR FILING DATE: 2000-07-12
NUMBER OF SEQ ID NOS: 407
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 328
LENGTH: 424
TYPE: PRT
ORGANISM: Eucalyptus grandis
US-10-174-693-328

Query Match 65.9%; Score 29; DB 16; Length 424;
Best Local Similarity 71.4%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQY 7
DB 101 NKGTDY 107

RESULT 32
US-10-002-784A-27
Sequence 27, Application US/10002784A
Publication No. US20030036644A1
GENERAL INFORMATION:
APPLICANT: Ulrich, Robert G.
APPLICANT: Ulrich, Robert G.
TITLE OF INVENTION: Bacterial Superantigen Vaccines
FILE REFERENCE: 003/233/SAP
CURRENT APPLICATION NUMBER: US/10/002,784A
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776
PRIOR FILING DATE: 97-06-25; 98-09-01
NUMBER OF SEQ ID NOS: 40

SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 27
LENGTH: 468
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: mutant Spea-mutant Speb fusion
US-10-002-784A-27

Query Match
Best Local Similarity 65.9%; Score 29; DB 15; Length 468;
Best Local Similarity 62.5%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
DB 283 NKGLKDYT 290

RESULT 33
US-021-425-28
Sequence 28, Application US/10021425
Publication No. US20030148420A1

GENERAL INFORMATION:
APPLICANT: Suzanne L. Bolten
APPLICANT: Alan M. Easton
APPLICANT: Leslie C. Engel
APPLICANT: Dean M. Messing
APPLICANT: John S. Ng
APPLICANT: Beverly A. Reitz
APPLICANT: Scott A. Vaccaro
APPLICANT: Mark C. Walker
APPLICANT: Ping T. Wang
APPLICANT: Robin A. Weinberg
TITLE OF INVENTION: Aspergillus ochraceus 11 alpha
TITLE OF INVENTION: Hydroxylase and oxidoreductase
FILE REFERENCE: S03196-00-US
CURRENT APPLICATION NUMBER: US/10/021,425
CURRENT FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: USSN 60/244,300
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 65
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 514
TYPE: PRT
ORGANISM: Neurospora crassa CAB91316
US-10-021-425-28

Query Match
Best Local Similarity 65.9%; Score 29; DB 12; Length 514;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYT 8
DB 211 KVTQOYT 217

RESULT 34
US-09-964-992A-5
Sequence 5, Application US/09964992A
Patent No. US20020173633A1
GENERAL INFORMATION:
APPLICANT: Salveit, Mikal E.
APPLICANT: Campos, Reinaldo
APPLICANT: No. US20020173633A1logaki, Hiroyuki
APPLICANT: Suslow, Trevor
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Characterization of Phenylalanine Ammonia-Lyase (PAL)
FILE REFERENCE: 023070-124500US
CURRENT APPLICATION NUMBER: US/09/964,992A
CURRENT FILING DATE: 2002-07-09
PRIOR APPLICATION NUMBER: US 60/235,956

PRIOR FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 5
LENGTH: 666
TYPE: PRT
ORGANISM: Helianthus annuus
FEATURE:
OTHER INFORMATION: sunflower phenylalanine ammonia-lyase (PAL)
US-09-964-992A-5

Query Match
Best Local Similarity 65.9%; Score 29; DB 10; Length 666;
Best Local Similarity 71.4%; Pred. No. 7.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 7
DB 97 NKGTDSY 103

RESULT 35
US-09-964-992A-1
Sequence 1, Application US/09964992A
Patent No. US20020173633A1
GENERAL INFORMATION:
APPLICANT: Salveit, Mikal E.
APPLICANT: Campos, Reinaldo
APPLICANT: No. US20020173633A1logaki, Hiroyuki
APPLICANT: Suslow, Trevor
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Characterization of Phenylalanine Ammonia-Lyase (PAL)
FILE REFERENCE: 023070-124500US
CURRENT APPLICATION NUMBER: US/09/964,992A
CURRENT FILING DATE: 2002-07-09
PRIOR APPLICATION NUMBER: US 60/235,956
PRIOR FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 711
TYPE: PRT
ORGANISM: Lactuca sativa
FEATURE:
OTHER INFORMATION: lettuce phenylalanine ammonia-lyase (PAL)* 1
US-09-964-992A-1

Query Match
Best Local Similarity 65.9%; Score 29; DB 10; Length 711;
Best Local Similarity 71.4%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 7
DB 97 NKGTDSY 103

RESULT 36
US-09-964-992A-2
Sequence 2, Application US/09964992A
Patent No. US20020173633A1
GENERAL INFORMATION:
APPLICANT: Salveit, Mikal E.
APPLICANT: Campos, Reinaldo
APPLICANT: No. US20020173633A1logaki, Hiroyuki
APPLICANT: Suslow, Trevor
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Characterization of Phenylalanine Ammonia-Lyase (PAL)
FILE REFERENCE: 023070-124500US
CURRENT APPLICATION NUMBER: US/09/964,992A
CURRENT FILING DATE: 2002-07-09
PRIOR APPLICATION NUMBER: US 60/235,956

;; PRIOR FILING DATE: 2000-09-26
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 713
;; TYPE: PRT
;; ORGANISM: Lactuca sativa
;; FEATURE:
;; OTHER INFORMATION: lettuce phenylalanine ammonia-lyase (PAL) 2
US-09-964-992A-2

Query Match 65.9%; Score 29; DB 10; Length 713;
Best Local Similarity 71.4%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOY 7
Db 102 NKGTDSY 108

ULT 37

US-09-854-122-49
;; Sequence 49, Application US/09854122
;; Patent No. US20020016980A1
;; GENERAL INFORMATION:
;; APPLICANT: ALBERTE, RANDALL S.
;; APPLICANT: SMITH, ROBERT
;; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
;; FILE REFERENCE: PHA-007.01
;; CURRENT APPLICATION NUMBER: US/09/854.122
;; PRIOR FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/202,529
;; PRIOR FILING DATE: 2000-05-10
;; NUMBER OF SEQ ID NOS: 51
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 49
;; LENGTH: 717
;; TYPE: PRT
;; ORGANISM: Arabidopsis thaliana
US-09-854-122-49

Query Match 65.9%; Score 29; DB 9; Length 717;
Best Local Similarity 71.4%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOY 7
Db 103 NKGTDSY 109

RESULT 38
US-10-218-102-103
;; Sequence 103, Application US/10218102
;; Publication No. US20030130827A1
;; GENERAL INFORMATION:
;; APPLICANT: Bentzien, Joerg
;; APPLICANT: Desjardais, Basell I.
;; APPLICANT: Desjardais, John R.
;; APPLICANT: Hayes, Robert J.
;; APPLICANT: Vielmetter, Jost
;; TITLE OF INVENTION: Protein Design Automation for Protein Libraries
;; FILE REFERENCE: A-67229-11/RFT/RMS/RMK
;; CURRENT APPLICATION NUMBER: US/10/218.102
;; PRIOR FILING DATE: 2002-08-12
;; PRIOR APPLICATION NUMBER: US 09/927,790
;; PRIOR FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: US 60/311,545
;; PRIOR FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: US 60/324,899
;; PRIOR FILING DATE: 2001-09-25
;; PRIOR APPLICATION NUMBER: US 60/351,937
;; PRIOR FILING DATE: 2002-01-25

;; PRIOR APPLICATION NUMBER: US 60/352,103
;; PRIOR FILING DATE: 2002-01-25
;; NUMBER OF SEQ ID NOS: 432
;; SOFTWARE: Patentin version 3.1
;; SEQ ID NO 103
;; LENGTH: 34
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: synthetic
US-10-218-102-103

Query Match 63.6%; Score 28; DB 16; Length 34;
Best Local Similarity 75.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
Db 21 NKDTQOYT 28

RESULT 39

US-10-218-102-226
;; Sequence 226, Application US/10218102
;; Publication No. US20030130827A1
;; GENERAL INFORMATION:
;; APPLICANT: Bentzien, Joerg
;; APPLICANT: Desjardais, Basell I.
;; APPLICANT: Desjardais, John R.
;; APPLICANT: Hayes, Robert J.
;; APPLICANT: Vielmetter, Jost
;; TITLE OF INVENTION: Protein Design Automation for Protein Libraries
;; FILE REFERENCE: A-67229-11/RFT/RMS/RMK
;; CURRENT APPLICATION NUMBER: US/10/218.102
;; PRIOR FILING DATE: 2002-08-12
;; PRIOR APPLICATION NUMBER: US 09/927,790
;; PRIOR FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: US 60/311,545
;; PRIOR FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: US 60/324,899
;; PRIOR FILING DATE: 2001-09-25
;; PRIOR APPLICATION NUMBER: US 60/351,937
;; PRIOR FILING DATE: 2002-01-25
;; PRIOR APPLICATION NUMBER: US 60/352,103
;; PRIOR FILING DATE: 2002-01-25
;; NUMBER OF SEQ ID NOS: 432
;; SOFTWARE: Patentin version 3.1
;; SEQ ID NO 226
;; LENGTH: 34
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: synthetic
US-10-218-102-226

Query Match 63.6%; Score 28; DB 16; Length 34;
Best Local Similarity 75.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
Db 21 NKDTQOYT 28

RESULT 40
US-10-218-102-238
;; Sequence 238, Application US/10218102
;; Publication No. US20030130827A1
;; GENERAL INFORMATION:
;; APPLICANT: Bentzien, Joerg
;; APPLICANT: Desjardais, Basell I.
;; APPLICANT: Desjardais, John R.
;; APPLICANT: Hayes, Robert J.

```
/ APPLICANT: Vielmeier, Josec
/ TITLE OF INVENTION: Protein Design Automation for Protein Libraries
/ FILE REFERENCE: A-67229-11/RFT/RMS/RMK
/ CURRENT APPLICATION NUMBER: US/10/218,102
/ CURRENT FILING DATE: 2002-08-12
/ PRIOR APPLICATION NUMBER: US 09/927,790
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: US 60/311,545
/ PRIOR FILING DATE: 2001-09-10
/ PRIOR APPLICATION NUMBER: US 60/324,899
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US 60/351,937
/ PRIOR FILING DATE: 2002-01-25
/ PRIOR APPLICATION NUMBER: US 60/352,103
/ PRIOR FILING DATE: 2002-01-25
/ NUMBER OF SEQ ID NOS: 432
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 238
/ LENGTH: 34
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: synthetic
US-10-218-102-238

Query Match          63.6%; Score 28; DB 16; Length 34;
Best Local Similarity 75.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 NKGTOQYT 8
        |||||
Db      21 NKDTQOST 28

Search completed: August 20, 2003, 10:08:17
Job time : 32.4948 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 8.9869 Seconds
(without alignments)
85.581 Million cell updates/sec

Title: US-09-991-433-6
Perfect score: 44
Sequence: 1 NKGTQOYT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

al number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	781	1 VCPV19	coat protein VP1 -
2	33	75.0	106	2 A89872	conserved hypotet
3	33	75.0	863	2 B63301	signal-transducing
4	32	72.7	163	2 T48609	hypothetical prote
5	32	72.7	240	2 S39047	trypsin (EC 3.4.21
6	32	72.7	290	2 C45240	cell shape determi
7	32	72.7	290	2 B68660	cell shape determi
8	32	72.7	290	2 UC4595	cell shape determi
9	32	72.7	362	2 T27272	hypothetical prote
10	31	70.5	1164	2 B71429	phytochrome D - Ar
11	31	70.5	204	2 A32252	probable phosphos
12	31	70.5	288	2 I51332	signal sequence re
13	31	70.5	1332	2 T23024	hypothetical prote
14	31	70.5	2163	2 S50675	pre-mRNA splicing
15	30	68.2	54	2 S31918	myrosinase - chick
16	30	68.2	60	1 WZVZ47	5K HindIII-C-prote
17	30	68.2	121	2 T29993	hypothetical prote
18	30	68.2	195	2 T36935	probable secreted
19	30	68.2	236	2 A28566	T-cell suppressor
20	30	68.2	241	2 S69131	Ig heavy chain (DO
21	30	68.2	258	4 S70439	pancreatic elastas
22	30	68.2	266	1 ELRT1	pancreatic elastas
23	30	68.2	266	1 ELRG	pancreatic elastas
24	30	68.2	267	4 A56615	probable pancreati
25	30	68.2	304	2 H75378	hypothetical prote
26	30	68.2	326	1 G2HU	Ig gamma-2 chain C
27	30	68.2	338	2 A53066	CCAAT enhancer-bin
28	30	68.2	360	2 T51870	hypothetical prote
29	30	68.2	366	2 F86794	hypothetical prote

30	30	68.2	398	2 D83956	hypothetical prote
31	30	68.2	505	2 C46642	DNA primase (EC 2.
32	30	68.2	509	2 S45631	DNA primase chain
33	30	68.2	540	2 T44683	precortin methylas
34	30	68.2	547	2 UC4519	heat-shock protein
35	30	68.2	548	2 C64076	secreted protein k
36	30	68.2	732	2 T43619	protein kinase Ypk
37	30	68.2	732	2 S30060	probable aldehyde
38	30	68.2	737	2 H95882	hypothetical prote
39	30	68.2	880	2 T04523	DNA dependent Atpa
40	30	68.2	1121	2 S30862	mycosubtilin synth
41	30	68.2	1335	2 AD2220	hypothetical prote
42	30	68.2	3971	2 T44806	hypothetical prote
43	29	65.9	144	2 A84126	tubulin beta chain
44	29	65.9	163	2 S39821	acetyl transferase
45	29	65.9	193	2 B86716	

ALIGNMENTS

RESULT 1
VCPV19
coat protein VP1 - parvovirus B19 (strain Au)
C:Species: parvovirus B19
A>Note: host Homo sapiens (man)
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C:Accession: A24299
R:Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.
J. Virol. 58, 921-936, 1986
A:Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr
A:Reference number: A24299; MUID:86200451; PMID:701931
A:Accession: A24299
A:Molecule type: DNA
A:Residues: 1-781 <SHA>
A:Cross-references: EMBL:M13178; NID:G333375; PIDN:AAA6867.1; PID:G333377
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 100.0%; Score 44; DB 1; Length 781;

Best Local Similarity 100.0%; Pred. No. 0.21; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Indels 0; Gaps 0;

Db 622 NKGTQOYT 629

RESULT 2
A89872
conserved hypothetical protein SA0887 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: A89872
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc

me, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: A89872
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-106 <KUR>
A:Cross-references: GB:BA000018; PID:gl3700836; PIDN:BA842132.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:

A:Gene: SA0887

Query Match 75.0%; Score 33; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NKGTQOY 7
 |||||
 Db 45 NKGTQOY 50

RESULT 3

signal-transducing histidine kinase homolog Af0410 - Archaeoglobus fulgidus
 C/Species: Archaeoglobus fulgidus
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
 C/Accession: B69301
 R/Name: H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A/Authors: Ueberlack, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaime, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A/Reference number: A69250; PMID:98049343; PMID:9389475
 C/Accession: B69301
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-863 <FUE>
 A/Cross-references: GB:AE001076; GB:AE000782; NID:92689399; PIDN:AA90824.1; PID:9265021

Query Match 75.0%; Score 33; DB 2; Length 863;
 Best Local Similarity 75.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
 |||||
 Db 433 NKATQOYT 440

RESULT 4

hypothetical protein F18022.150 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cross)
 C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C/Accession: T48609
 R/Name: M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, April 2000
 A/Reference number: Z24493
 A/Accession: T48609
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-163 <BEV>
 A/Cross-references: EMBL:DB
 A/Note: experimental source: cultivar Columbia; BAC clone F18022
 A/Map position: 5
 A/Intons: 92/1
 A/Note: F18022.150

Query Match 72.7%; Score 32; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQO 6
 |||||
 Db 21 NKGTQO 26

RESULT 5

tryptase (EC 3.4.21.4) I - Atlantic cod
 C/Species: Gadus morhua (Atlantic cod)
 C/Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
 C/Accession: S39047
 R/Name: Gudmundsdottir, A.; Gudmundsdottir, E.; Oskarsson, S.; Bjarnason, J.B.; Eakin, A.K.; C
 Eur. J. Biochem. 217, 1091-1097, 1993
 A/Title: Isolation and characterization of cDNAs from Atlantic cod encoding two different
 A/Reference number: S39047; PMID:94039130; PMID:8223632

A/Accession: S39047
 A/Molecule type: mRNA
 A/Residues: 1-240 <GUD>
 A/Cross-references: EMBL:X76886; NID:9450517; PIDN:CAA54214.1; PID:91334752
 C/Suprafamily: trypsin; trypsin homology
 C/Keywords: hydrolase; serine proteinase
 F/19-233/Domain: trypsin homology <TRY>
 F/58,102,194/Active site: His, Asp, Ser #status predicted

Query Match 72.7%; Score 32; DB 2; Length 240;
 Best Local Similarity 71.4%; Pred. No. 25;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
 |||||
 Db 76 NKGTQOY 82

RESULT 6

cell shape determinant mreC - Bacillus subtilis
 C/Species: Bacillus subtilis
 C/Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
 C/Accession: C45240; D45239; S27518
 R/Name: Varley, A.W.; Stewart, G.C.
 J. Bacteriol. 174, 6729-6742, 1992
 A/Title: The divIB region of the Bacillus subtilis chromosome encodes homologs of Esche
 A/Reference number: A45240; PMID:93015732; PMID:1400225
 A/Accession: C45240
 A/Molecule type: DNA
 A/Residues: 1-290 <VAR>
 A/Cross-references: EMBL:M95582; NID:9143211; PIDN:AAA22606.1; PID:9143213
 A/Note: sequence extracted from NCBI backbone (NCBI:116571)
 R/Name: Varley, A.W.; Margolis, P.S.; Setlow, P.; Loeb, R.; Sun, D.
 J. Bacteriol. 174, 6717-6728, 1992
 A/Title: Identification of Bacillus subtilis genes for septum placement and shape determ
 A/Reference number: A45239; PMID:93015731; PMID:1400224
 A/Accession: D45239
 A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-142, 'P', 144-175, 'P', 177-290 <LEV>
 A/Cross-references: GB:M95343; NID:9142852; PIDN:AAA22398.1; PID:9142856
 A/Note: sequence extracted from NCBI backbone (NCBI:116559)
 C/Genetics:
 A/Genes: mreC

Query Match 72.7%; Score 32; DB 2; Length 290;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQO 6
 |||||
 Db 134 NKGTQO 139

RESULT 7

cell-shape determining protein mreC - Bacillus subtilis
 C/Species: Bacillus subtilis
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C/Accession: B69660
 R/Name: Kunet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Berte
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Cho
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
 Nature 390, 249-256, 1997
 A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Galizzi, A.; Galler
 koetter, J.; Harwood, C.R.; Hentut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mancel
 Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolet, C.; Rocha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanlon,
 A/Authors: Schleich, C.; Schroefer, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
 akouchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: B69660
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-290 <KUN>
 A:Cross-references: GB:299118; GB:AL009126; NID:G2635200; PIDN:CAB14762.1; PID:G2635267
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: mrec

Query Match 72.7%; Score 32; DB 2; Length 290;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQO 6
 |||||
 DB 134 NKGTQO 139

RESULT 8

JC4595
 cell shape determinant Mrec - *Bacillus stearothermophilus*
 M:Alternate names: protease secretion stimulating protein, Mrec
 C:Species: *Bacillus stearothermophilus*
 C:Date: 16-Apr-1996 #sequence_revision 24-May-1996 #text_change 26-Jul-1996
 C:Accession: JC4595
 R:Kubo, M.; Pietro, D.J.; Mochizuki, Y.; Kojima, T.; Yamazaki, T.; Satoh, S.; Takizawa, B.
 Biosci. Biotechnol. Biochem. 60, 271-276, 1996
 A:Title: *Bacillus stearothermophilus* cell shape determinant gene, mrec and mred, and the
 A:Reference number: JC4595; MUID:97076621; PMID:9063975
 A:Accession: JC4595
 A:Molecule type: DNA
 A:Residues: 1-290 <KUB>
 A:Experimental source: HAI9
 C:Comment: The gene of cell shape determinant regulates protease expression and stimulates
 C:Genetics:
 A:Gene: mrec

Query Match 72.7%; Score 32; DB 2; Length 290;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQO 6
 |||||
 DB 134 NKGTQO 139

RESULT 9

727272
 hypothetical protein Y63D3A.4 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27272
 R:White, S.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: Z20335
 A:Accession: T27272
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-362 <WIL>
 A:Cross-references: EMBL:AL032652; PIDN:CAA21707.1; GSPDB:GN00019; CESP:Y63D3A.4
 A:Experimental source: clone Y63D3A
 C:Genetics:
 A:Gene: CESP:Y63D3A.4
 A:Map position: 1
 A:Introns: 18/1; 43/1; 116/3; 148/2; 225/3; 320/2
 Query Match 72.7%; Score 32; DB 2; Length 362;
 Best Local Similarity 75.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
 |||||
 DB 180 NKGTQOYT 187

RESULT 10

B71429
 phytochrome D - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 A:Variety: Columbia
 C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
 C:Accession: B71429; S46312; S41910
 R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
 vanagh, T.; Hempel, S.; Kotter, P.; Ertan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
 Nature 391, 485-488, 1998
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdemecch
 erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
 C.; Chalwatzis, N.
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thal*
 A:Reference number: A71400; MUID:98121113; PMID:9461215
 A:Accession: B71429
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1164 <BEV>
 A:Cross-references: GB:297340; NID:G2244950; PIDN:CAB10404.1; PID:G2244983
 R:Clack, T.; Mathews, S.; Sharrock, R.A.
 Plant Mol. Biol. 25, 413-427, 1994
 A:Title: The phytochrome apoprotein family in *Arabidopsis* is encoded by five genes: the
 A:Reference number: S46312; MUID:94325466; PMID:8049367
 A:Accession: S46312
 A:Molecule type: DNA
 A:Residues: 1-424, 'F', 426-1164 <CLA>
 A:Cross-references: EMBL:X76609; NID:9452812; PIDN:CA54072.1; PID:9452814
 C:Genetics:
 A:Gene: PHVD
 A:Map position: 4COP9-4G3845
 C:Superfamily: phytochrome; phytochrome homology
 C:Keywords: chromoprotein; photoreceptor; phytochromobilin; transcription regulation
 F:103-618/Domain: phytochrome homology <PHYT>
 F:360/Binding site: phytochromobilin (Cys) (covalent) #status predicted

Query Match 72.7%; Score 32; DB 2; Length 1164;
 Best Local Similarity 75.0%; Pred. No. 1,4e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
 |||||
 DB 57 NKGTQOYT 64

RESULT 11
 A32252
 probable phosphoesterase (EC 3.1.-.-) L - *Acinetobacter calcoaceticus* (fragment)
 N:Alternate names: gene L protein
 C:Species: *Acinetobacter calcoaceticus*
 C:Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 24-Sep-1999
 C:Accession: A32252; T01631
 R:Goosen, N.; Horman, H.P.A.; Hulnen, R.G.M.; van de Putte, P.
 J. Bacteriol. 171, 447-455, 1989
 A:Title: *Acinetobacter calcoaceticus* genes involved in biosynthesis of the coenzyme pyr.
 A:Reference number: A32252; MUID:89123056; PMID:2536663
 A:Accession: A32252
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-204 <GOO>
 A:Cross-references: GB:X06452; NID:G38740; PIDN:CAA29752.1; PID:G38741
 A:Experimental source: strain LMD 79.41
 C:Comment: This sequence has motifs characteristic of a variety of phosphoesterases.
 C:Superfamily: unassigned probable phosphoesterases; phosphoesterase core homology
 C:Keywords: hydrolase
 F:43-139/Domain: phosphoesterase core homology <PEC>

Query Match 70.5%; Score 31; DB 2; Length 204;
 Best Local Similarity 75.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
 |||||
 DB 152 NGSCQOYT 159

RESULT 12

151332
 signal sequence receptor alpha chain - rainbow trout
 C:Species: Oncorhynchus mykiss (rainbow trout)
 C>Date: 13-Sep-1996 #sequence_rev151332
 C:Accession: 151332
 R:Hartmann, E.; Prehn, S.
 FEBS Lett. 349, 324-326, 1994
 A>Title: The N-terminal region of the alpha-subunit of the TRAP complex has a conserved
 A:Reference number: 138246; MID:94326944; PMID:8050590
 A:Accession: 151332
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-288 <HAR>
 A:Cross-references: EMBL:Z12831; NID:G541750; PIDN:CAA78291.1; PID:G541751
 C:Superfamily: human signal sequence receptor alpha chain

Query Match 70.5%; Score 31; DB 2; Length 288;
 Best Local Similarity 62.5%; Pred. No. 51;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
 |||||
 DB 110 NKGSQDPT 117

RESULT 13

T23024
 hypothetical protein H02112.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_rev151332
 C:Accession: T23024
 R:McLay, K.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19656
 A:Accession: T23024
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1332 <MIL>
 A:Cross-references: EMBL:Z92769; PIDN:CAB07215.1; GSPDB:GN00022; CESP:H02112.1
 A:Experimental source: clone H02112
 C:Genetics:
 A:Gene: CESP:H02112.1
 A:Map position: 4
 A:Introns: 12/3; 29/1; 60/3; 246/1; 449/1; 756/1; 849/1; 961/1; 1103/3; 1214/3

Query Match 70.5%; Score 31; DB 2; Length 1332;
 Best Local Similarity 75.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
 |||||
 DB 579 NKDCQOYT 586

RESULT 14

S50675
 pre-mRNA splicing helicase BRR2 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: DNA repair protein RAD24; protein YER172c
 C:Species: Saccharomyces cerevisiae
 C>Date: 28-May-1993 #sequence_rev151332
 C:Accession: S50675; S30856; J04657
 R:Dietrich, F.S.

submitted to the EMBL Data Library, December 1994
 A:Description: The sequence of S. cerevisiae cosmid 9163 and 9132.
 A:Reference number: S50431
 A:Accession: S50675

A:Molecule type: DNA
 A:Residues: 1-2163 <DIE>
 A:Cross-references: EMBL:U18922; NID:G603405; PIDN:AAB64699.1; PID:G603413; MIPS:YER172c
 R:Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor,
 submitted to the EMBL Data Library, February 1993
 A:Reference number: S30812
 A:Accession: S30856

A:Molecule type: DNA
 A:Residues: 1-169 <MUL>
 A:Cross-references: EMBL:L11229
 R:Zhu, Y.B.; Prakash, L.; Prakash, S.
 Chinese Biochem. J. 11, 541-550, 1995
 A>Title: Molecular cloning and sequencing of DNA repair gene RAD24.
 A:Reference number: J04657
 A:Accession: J04657
 A:Molecule type: DNA
 A:Residues: 1-260, 'HCOT', 262, 'KT', 265 <ZHU>

C:Genetics:
 A:Gene: SGD:BRR2; RAD24
 A:Cross-references: SGD:S000974; MIPS:YER172c
 A:Map position: 5R
 C:Keywords: ATP; DNA repair; nucleotide binding; nucleus; P-loop; pre-mRNA splicing
 F:521-528/Region: nucleotide-binding motif A (P-loop)
 F:630-635/Region: nucleotide-binding motif B
 F:634-637/Region: DEXH motif

Query Match 70.5%; Score 31; DB 2; Length 2163;
 Best Local Similarity 85.7%; Pred. No. 4.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 7
 |||||
 DB 2120 NKETQOYT 2126

RESULT 15

S31918
 myrosinase - chickpea (fragment)
 C:Species: Cicer arctium (chickpea, garbanzo)
 C>Date: 06-Feb-1995 #sequence_rev151332
 C:Accession: S31918
 R:Cervantes, E.
 submitted to the EMBL Data Library, February 1993
 A:Reference number: S31914
 A:Accession: S31918
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-54 <CER>
 A:Cross-references: EMBL:X70374; NID:G22672; PIDN:CAA49835.1; PID:G536768
 C:Superfamily: Agrobacterium beta-glucosidase

Query Match 68.2%; Score 30; DB 2; Length 54;
 Best Local Similarity 62.5%; Pred. No. 14;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
 |||||
 DB 15 NSGTETPT 22

RESULT 16

WZVZ47
 5K HindIII-C protein - vaccinia virus (strain WR)
 C:Species: vaccinia virus
 C>Date: 31-Mar-1990 #sequence_rev151332
 C:Accession: G31829
 R:Kolwal, G.J.; Moss, B.
 Virology 167, 524-537, 1988
 A>Title: Analysis of a large cluster of nonessential genes deleted from a vaccinia virus

A:Reference number: A94385; MUID:89073756; PMID:2849238
 A:Accession: G31829
 A:Molecule type: DNA
 A:Residues: 1-60 <NOT>
 A:Cross-references: GB:M22812; NID:G315691; PIDN:AA95958.1; PID:G893307
 C:Superfamily: vaccinia virus 5K HindIII-C protein
 C:Keywords: early protein

Query Match 68.2%; Score 30; DB 1; Length 60;
 Best Local Similarity 62.5%; Pred. No. 16;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
 |||:|
 Db 23 NKGSRTYT 30

RESULT 17

T29993
 Spectral protein C43H6.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C:Accession: T29993

R:File: T.T.
 Submitted to the EMBL Data Library, March 1996
 A:Description: The sequence of C. elegans cosmid C43H6.
 A:Reference number: Z20717
 A:Accession: T29993

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-121 <LET>
 A:Cross-references: EMBL:U51999; PIDN:AA96087.1; GSPDB:GN00028; CESP:C43H6.2
 A:Experimental source: strain Bristol N2; clone C43H6

C:Genetics:
 A:Gene: CESP:C43H6.2
 A:Map position: X
 A:Introns: 22/3; 53/3; 76/3

C:Superfamily: Caenorhabditis elegans hypothetical protein C43H6.2

Query Match 68.2%; Score 30; DB 2; Length 121;
 Best Local Similarity 71.4%; Pred. No. 33;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTOOYT 8
 |||:|
 Db 29 RGTMOYT 35

T18
 35
 Probable secreted protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T18935
 R:Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, September 1999

A:Reference number: Z21607
 A:Accession: T36935

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-195 <SEB>
 A:Cross-references: EMBL:AL109962; PIDN:CAB53120.1; GSPDB:GN00070; SCOEDB:SCJ1.02c
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SCJ1.02c

Query Match 68.2%; Score 30; DB 2; Length 195;
 Best Local Similarity 71.4%; Pred. No. 55;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYT 7
 |||:|
 Db 156 NKGSQNTY 162

RESULT 19

A28566
 T-cell suppressor factor, phosphorylcholine-specific - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 15-Dec-1988 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999

C:Accession: A28566
 R:Yamasaki, N.; Sugimura, K.; Hida, M.; Naito, T.; Watanabe, T.
 Eur. J. Immunol. 17, 247-253, 1987

A:Title: Sequence analysis of a cDNA clone of a gene encoding a component of a putative
 A:Reference number: A28566; MUID:87162073; PMID:343908

A:Accession: A28566
 A:Molecule type: mRNA
 A:Residues: 1-236 <YAM>

A:Cross-references: GB:M27347; NID:G200267; PIDN:AA39901.1; PID:G200268
 A:Experimental source: T-cell hybridoma

A>Note: the authors translated the codon AAC for residue 69 as Val, GTG for residue 74 as His, and CTC for residue 222 as Val
 C:Superfamily: trypsin; trypsin homology
 F:1-229/Domain: trypsin homology (fragment) <TRY>

Query Match 68.2%; Score 30; DB 2; Length 236;
 Best Local Similarity 71.4%; Pred. No. 68;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYT 7
 |||:|
 Db 61 NDGTQOYT 67

RESULT 20

S69131
 Ig heavy chain (DOT) - human (fragment)
 N:Alternate names: anti-riboflavin IgG Fd fragment

C:Species: Homo sapiens (man)
 C>Date: 12-Feb-1998 #sequence_revision 22-May-1998 #text_change 21-Jan-2000

C:Accession: S69131
 R:Stoppini, M.; Bellotti, V.; Negri, A.; Merlini, G.; Garver, F.; Ferri, G.
 Eur. J. Biochem. 228, 886-893, 1995

A:Title: Characterization of the two unique human anti-flavin monoclonal immunoglobulins
 A:Reference number: S69130; MUID:9525298; PMID:7737190

A:Accession: S69131
 A:Molecule type: protein
 A:Residues: 1-241 <STO>

C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: blocked amino end; heterotetramer; immunoglobulin; pyroglyutamic acid
 F:1-241/Product: Ig heavy chain (DOT) (fragment) #status experimental <MAT>
 F:140-205/Domain: immunoglobulin homology <IMM>
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 68.2%; Score 30; DB 2; Length 241;
 Best Local Similarity 75.0%; Pred. No. 69;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
 |||:|
 Db 195 NFGTQOYT 202

RESULT 21

S70439
 pancreatic elastase I (allele HBL1-16) probable splice form I - human
 C:Species: Homo sapiens (man)

C>Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 28-Apr-2003
 C:Accession: S70439

R:Kawashima, I.; Tani, T.; Mita-Honjo, K.; Shimoda-Takano, K.; Ohmine, T.; Furukawa, H.;
 DNA Seq. 2, 303-312, 1992
 A:Title: Genomic organization of the human homologue of the rat pancreatic elastase I ge

A:Reference number: A56615; MUID:9233895; PMID:1633328
 A:Accession: S70439
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

A:Residues: 1-258 <RAW>
A>Note: It is not known whether the gene is expressed
F:19-251/Domain: trypsin homology <TRY>

Query Match 68.2%; Score 30; DB 4; Length 258;
Best Local Similarity 71.4%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
DB 83 NDGTQOY 89

RESULT 22

ELR1

pancreatic elastase (EC 3.4.21.36) I precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 18-Jun-1999

C:Accession: A00960; A20534

C:Keywords: R.U.; Swift, G.H.; Quinto, C.; Swain, W.; Pictet, R.L.; Nikovits, W.; Rutledge

Chemistry 21, 1453-1463, 1982

A>Title: Primary structure of two distinct rat pancreatic preproelastases determined by

A:Reference number: A00960; PMID:82182967; PMID:6918221

A:Accession: A00960

A:Molecule type: mRNA

A:Residues: 1-266 <MAC>

A:Cross-references: GB:V01234; NID:G56088; PIDN:CAA24544.1; PID:G56089

R:Jargman, C.

Biochemistry 22, 3763-3770, 1983

A>Title: Isolation and characterization of rat pancreatic elastase.

A:Reference number: A20534; PMID:84000385; PMID:6555050

A:Accession: A20534

A:Molecule type: protein

A:Residues: 17-37,'X',39-45 <LAR>

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; serine proteinase; zymogen

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-26/Domain: activation peptide #status predicted <APR>

F:27-266/Product: elastase I #status predicted <MPT>

F:27-259/Domain: trypsin homology <TRY>

F:11,119,214/Active site: His, Asp, Ser #status predicted

Query Match 68.2%; Score 30; DB 1; Length 266;
Best Local Similarity 71.4%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
DB 91 NDGTQOY 97

RESULT 23

ELPG

pancreatic elastase (EC 3.4.21.36) I precursor - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 24-Apr-1984 #sequence_revision 30-Sep-1990 #text_change 16-Jun-2000

C:Accession: J00013; A26777; A10061; A00959

C:Keywords: Y.; Yoshida, H.; Miyakawa, T.; Matsuki, S.; Tanaka, J.I.; Ikenaga, H.

J. Biochem. 99, 1707-1712, 1986

A>Title: Isolation and expression in Escherichia coli of a cDNA clone encoding porcine P

A:Reference number: A92005; PMID:86304235; PMID:3528137

A:Accession: J00013

A:Molecule type: mRNA

A:Residues: 1-266 <SHI>

A:Cross-references: GB:X04036; GB:D00070; GB:N00070; NID:G1941; PIDN:CAA27670.1; PID:G19

R:Tani, T.; Kawashima, I.; Furukawa, H.; Ohmine, T.; Takiguchi, Y.

J. Biochem. 101, 591-599, 1987

A>Title: Characterization of a cDNA clone for human pancreatic elastase I: structure of

A:Accession: A26777

A:Molecule type: mRNA

A:Residues: 1-125,'G',127-183,'L',185-266 <TAN>

A:Cross-references: GB:D00160; NID:G217683; PIDN:BA00118.1; PID:G217684

A>Note: the authors translated the codon GCG for residue 58 as Gln, GGC for residue 126
R:Shotton, D.M.; Hartley, B.S.
Biochem. J. 131, 643-675, 1973

A>Title: Evidence for the amino acid sequence of porcine pancreatic elastase.

A:Reference number: A90267; PMID:73292121; PMID:4578945

A:Accession: A10061

A:Molecule type: protein

A:Residues: 27-91,'N',93-203,'N',205-266 <SHO>

R:Shotton, D.M.; Hartley, B.S.

Nature 225, 811-816, 1970

A>Title: Three-dimensional structure of tosyl-elastase.

A:Reference number: A91160; PMID:70114044; PMID:5415110

A:Contents: annotation; X-ray crystallography, 3.5 angstroms; active site

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; serine proteinase; zymogen

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-26/Domain: activation peptide #status predicted <APR>

F:27-266/Product: elastase I #status experimental <MAT>

F:56-72,153-220,184-200,210-240/Disulfide bonds: #status experimental

F:11,119,214/Active site: His, Asp, Ser #status experimental

Query Match 68.2%; Score 30; DB 1; Length 266;
Best Local Similarity 71.4%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
DB 91 NDGTQOY 97

RESULT 24

A56615

probable pancreatic elastase (EC 3.4.21.36) pseudogene - human

N:Alternate names: pancreatic elastase I homolog; pancreatic elastase I allele HEL1-16,

C:Species: Homo sapiens (man)

C>Date: 11-Aug-1995 #sequence_revision 17-Aug-1995 #text_change 14-Aug-1998

C:Accession: A56615; S70440

R:Kawashima, I.; Tani, T.; Mita-Honjo, K.; Shimoda-Takano, K.; Ohmine, T.; Furukawa, H.,

DNA Seq. 2, 303-312, 1992

A>Title: Genomic organization of the human homologue of the rat pancreatic elastase I ge

A:Reference number: A56615; PMID:92338395; PMID:1633328

A:Accession: A56615

A:Molecule type: DNA

A:Residues: 1-267 <RAW>

A:Cross-references: EMBL:X62259; NID:G31246; EMBL:X62258; GB:S40923; NID:G31247; EMBL:X6

5; GB:S40856; NID:G31251; EMBL:X62256; GB:S40857; NID:G31252; EMBL:X62257; GB:S40859; NI

A>Note: sequence extracted from NCBI backbone (NCBI:109315, NCBI:109319, NCBI:109319,

C:Comment: This apparently silent human homolog of pancreatic elastase I is a single-cop

unctional protein in some other tissue.

C:Genetics:

A:Gene: GDB:ELAI

A:Map position: 12

C:Keywords: hydrolase; pseudogene; serine proteinase

Query Match 68.2%; Score 30; DB 4; Length 267;
Best Local Similarity 71.4%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
DB 92 NDGTQOY 98

RESULT 25

H75378

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000

C:Accession: H75378

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.U.;

M.; Shen, M.; Vanatrevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999
 A>Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A'Reference number: A75250; MUID:20036896; PMID:10567266
 A'Accession: H75378
 A'Molecule type: Preliminary
 A'Results: 1-304 <WH>
 A'CROSS-references: GB:AE002002; GB:AE000513; NID:g6459345; PIDN:AAFI1154.1; PID:g645935
 A'Experimental source: strain R1
 C'Genetics:
 A'Gene: DR1591
 A'Map position: 1
 C'Superfamily: *Deinococcus radiodurans* hypothetical protein DR1591

Query Match 68.2% Score 30; DB 2; Length 304;
 Best Local Similarity 62.5%; Pred. No. 89;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYT 8
 |||||
 286 NKGLENYT 293

RESULT 26
 G2HU
 I9 gamma-2 chain C region - human
 C'Species: *Homo sapiens* (man)
 C'Date: 30-Apr-1981 #sequence revision 13-Jun-1983 #text_change 21-Jul-2000
 C'Accession: A93906; A92809; A90752; A93132; A02148
 R'Ellison, J.; Hood, L.
 Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982
 A>Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con
 A'Reference number: A93906; MUID:82197621; PMID:6804948
 A'Accession: A93906
 A'Molecule type: DNA
 A'Results: 1-326 <EL>
 A'CROSS-references: GB:V00554; GB:J00230; NID:G33759; PIDN:CAB58438.1; PID:G6066056
 A'Note: Lys-326 is probably removed posttranslationally
 R'Wang, A.C.; Tung, E.; Pridenbery, H.H.
 J. Immunol. 125, 1048-1054, 1980
 A>Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f
 A'Reference number: A92809; MUID:81007873; PMID:6774012
 A'Accession: A92809
 A'Molecule type: Protein
 A'Results: 1-19, 'Q', 21-5', 'Z', 59, 'A', 61-193, 'D', 195-325 <MAN>
 A'Note: Tyr-156 is at or near the complement-binding site
 R'Connell, G.E.; Parr, D.M.; Hofmann, T.
 J. Biochem. 57, 758-767, 1979
 A>Title: The amino acid sequences of the three heavy chain constant region domains of a
 A'Reference number: A90752; MUID:80001357; PMID:113060
 A'Accession: A90752
 A'Molecule type: Protein
 A'Results: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-
 A'Note: this sequence has since been revised
 R'Hofmann, T.; Parr, D.M.
 Mol. Immunol. 16, 923-925, 1979
 A>Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g
 A'Reference number: A93132; MUID:80114419; PMID:118920
 A'Accession: A93132
 A'Molecule type: Protein
 A'Results: 238-275 <HOP>
 R'Hofmann, T.; Parr, D.M.
 Submitted to the Atlas, March 1980
 A'Reference number: A94591
 A'Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268
 A'Note: the revised sequence differs from that shown in having 60-Ala and in the amidati
 ned
 R'Milstein, C.; Frangione, B.
 Biochem. J. 121, 217-225, 1971
 A>Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.

A'Reference number: A90253; MUID:72033500; PMID:4940472
 A'Contents: annotation; myeloma protein Sa, disulfide bonds
 R'Frangione, B.; Milstein, C.; Pink, J.R.L.
 Nature 221, 145-148, 1969
 A>Title: Structural studies of immunoglobulin G.
 A'Reference number: A93157; MUID:69064124; PMID:5782707
 A'Contents: annotation; Sa, disulfide bonds
 C'Genetics:
 A'Gene: GDB:IGHG2
 A'CROSS-references: GDB:119338; OMTM:147110
 A'Map position: 14q32.33-14q32.33
 C'Complex: An immunoglobulin heterotetramer subunit consists of two identical light (K) f
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1s
 C'Superfamily: immunoglobulin C region; immunoglobulin homology
 C'Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F'20-85/Domain: immunoglobulin homology <IM1>
 F'113-202/Domain: immunoglobulin homology <IM2>
 F'239-306/Domain: immunoglobulin homology <IM3>
 F'14/Diulfide bonds: interchain (to light chain) #status experimental
 F'27-83, 140-200, 246-304/Diulfide bonds: #status experimental
 F'102, 103, 106, 109/Diulfide bonds: interchain (to heavy chain) #status experimental
 F'176/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 68.2% Score 30; DB 1; Length 326;
 Best Local Similarity 75.0%; Pred. No. 96;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYT 8
 |||||
 75 NFGTQTYT 82

RESULT 27
 A53066
 CCAAT enhancer-binding protein - California sea hare
 N'Alternate names: C/EBP
 C'Species: *Aplysia californica* (California sea hare)
 C'Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text_change 21-Jul-2000
 C'Accession: A53066
 R'Alberini, C.M.; Ghirardi, M.; Metz, R.; Kandel, E.R.
 Cell 76, 1099-1114, 1994
 A>Title: C/EBP is an immediate-early gene required for the consolidation of long-term fa
 A'Reference number: A53066; MUID:94185169; PMID:8137425
 A'Accession: A53066
 A'Status: preliminary
 A'Molecule type: mRNA
 A'Results: 1-338 <ALB>
 A'CROSS-references: GB:U00994; NID:G392976; PIDN:AA18286.1; PID:G487963

Query Match 68.2% Score 30; DB 2; Length 338;
 Best Local Similarity 83.3%; Pred. No. 99;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KGTQYT 7
 |||||
 258 KGTQYT 263

RESULT 28
 T51870
 hypothetical protein DKFZP547H084.1 - human
 C'Species: *Homo sapiens* (man)
 C'Date: 20-Oct-2000 #sequence revision 20-Oct-2000 #text_change 20-Oct-2000
 C'Accession: T51870
 R'Blocker, H.; Bocher, M.; Brandt, P.; Mewes, H.W.; Well, B.; Wiemann, S.
 submitted to the Protein Sequence Database, July 2000
 A'Reference number: Z25853
 A'Accession: T51870
 A'Status: preliminary
 A'Molecule type: mRNA
 A'Results: 1-360 <AAA>
 A'CROSS-references: EMBL:AL390139
 A'Experimental source: fetal brain; clone DKFZP547H084

C:Genetics:
A:Note: DKFZps47H084.1

Query Match

Best Local Similarity 68.2%; Score 30; DB 2; Length 360;
Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
|:|:|:|:
DB 307 NCKGKQKT 314

RESULT 29

D83794

hypothetical protein ynjG [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)

C:Species: *Lactococcus lactis* subsp. *lactis*
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: F86794

R:Polocin, A.; Mueker, P.; Mueger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich, J. M. *Proc. Natl. Acad. Sci. USA* 98:11731-11735, 2001

A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s8
A:Reference number: A86625; MUID:2125156; PMID:11337471

A:Accession: F86794

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-366 <STO>

A:Cross-references: GB:AE005176; PID:912724341; PIDN:AAK05456.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:
A:Gene: ynjG

Query Match

Best Local Similarity 68.2%; Score 30; DB 2; Length 366;
Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
|:|:|:|:
DB 353 NAGTFOYT 360

RESULT 30

D83956

hypothetical protein BH2452 [imported] - *Bacillus halodurans* (strain C-125)

C:Species: *Bacillus halodurans*

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: D83956

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai, T. *Proc. Natl. Acad. Sci. USA* 98:4317-4321, 2001

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: D83956

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-398 <STO>

A:Cross-references: GB:AF001515; GB:BA000004; NID:910174886; PIDN:BAB06171.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:
A:Gene: BH2452

Query Match

Best Local Similarity 68.2%; Score 30; DB 2; Length 398;
Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
|:|:|:|:
DB 193 NCKGKQKT 200

RESULT 31

C46642

DNA primase (EC 2.7.7.-) 54K chain - mouse

N/Alternate names: DNA polymerase alpha/DNA primase complex 54K chain

C:Species: *Mus musculus* (house mouse)

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000

C:Accession: C46642; S45629

R:Minazawa, H.; Izumi, M.; Tada, S.; Takada, R.; Maatani, M.; Ui, M.; Hanaka, F.

J. Biol. Chem. 268, 8111-8122, 1993

A:Title: Molecular cloning of the cDNAs for the four subunits of mouse DNA polymerase α

A:Reference number: A46642; MUID:93216788; PMID:8463324

A:Accession: C46642

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-505 <MTV>

A:Cross-references: GB:D13545; NID:9303662; PIDN:BA02745.1; PID:9303663

A:Experimental source: FMA cells

A:Note: sequence extracted from NCBI backbone (NCBI:129150, NCBI:129151)

R:Stadlbauer, F.; Brueckner, A.; Reifues, C.; Eckerstorn, C.; Lottspeich, F.; Foerster, E. *J. Biochem.* 222, 781-793, 1994

A:Title: DNA replication in vitro by recombinant DNA-polymerase- α -primase.

A:Reference number: S45628; MUID:94298818; PMID:8026492

A:Accession: S45629

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-108 'G', 110-505 <STV>

A:Cross-references: EMBL:D17385; NID:9397830; PIDN:BA04203.1; PID:9442472

C:Superfamily: mouse DNA primase 54K chain

C:Keywords: DNA binding; nucleotidyltransferase

Query Match

Best Local Similarity 68.2%; Score 30; DB 2; Length 505;
Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KGTQOY 7
|:|:|:|:
DB 65 KGTQOY 70

RESULT 32

S45631

DNA primase chain p58 - human

C:Species: *Homo sapiens* (man)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000

C:Accession: S45631

R:Stadlbauer, F.; Brueckner, A.; Reifues, C.; Eckerstorn, C.; Lottspeich, F.; Foerster, E. *J. Biochem.* 222, 781-793, 1994

A:Title: DNA replication in vitro by recombinant DNA-polymerase- α -primase.

A:Reference number: S45628; MUID:94298818; PMID:8026492

A:Accession: S45631

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-509 <STA>

A:Cross-references: EMBL:X74331; NID:9510407; PIDN:CAA52378.1; PID:9510408

C:Superfamily: mouse DNA primase 54K chain

Query Match

Best Local Similarity 68.2%; Score 30; DB 2; Length 509;
Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KGTQOY 7
|:|:|:|:
DB 65 KGTQOY 70

RESULT 33

T44683

precocorrin methylase (EC 2.1.1.-) [imported] - *Bacillus megaterium*

C:Species: *Bacillus megaterium*

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T44683

R:Raux, E.; Lanois, A.; Warren, M.J.; Rambach, A.; Thernes, C.

Biochem. J. 335, 159-166, 1998

A:Title: Cobalamin (vitamin B12) biosynthesis: identification and characterization of a

A:Reference number: Z22829; MUID:98416126; PMID:9742225

A:Accession: T44683

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Residues: 1-540 <RAU>
A:Cross-references: EMBL:AJ000758; NID:g3724036; PIDN:CAA04307.1; PID:g3724038
A:Experimental source: strain DSM 509
C:Genetics:
A:Note: cbih60
C:Keywords: methyltransferase

Query Match 68.2%; Score 30; DB 2; Length 540;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYT 8
|||:|
Db 356 KGTMEYT 362

RESULT 34

JC4519
heat-shock protein GroEL - Pasteurella multocida

A:Species: Pasteurella multocida
Accession: JC4519
Date: 08-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
R:Love, B.C.; Hansen, L.M.; Hirsch, D.C.
Gene 166, 179-180, 1995

A:Title: Cloning and sequence of the groEL heat-shock operon of Pasteurella multocida.
A:Reference number: JC4518; MUID:96105224; PMID:8529887
A:Accession: JC4519

A:Molecule type: DNA

A:Residues: 1-547 <LOV>
A:Cross-references: GB:U30165; NID:g1144300; PIDN:AAA84916.1; PID:g1144302

A:Experimental source: serotype A:3
C:Genetics:

A:Gene: groEL

C:Superfamily: chaperonin groEL

C:Keywords: heat shock; stress-induced protein

Query Match 68.2%; Score 30; DB 2; Length 547;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
|||:|
Db 479 NAGTEOY 485

RESULT 35

C64076
chaperonin groEL - Haemophilus influenzae (strain Rd KW20)

A:Species: Haemophilus influenzae
Accession: C64076
Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 26-Aug-1999

R:Flitschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.; Glodok, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Goughnagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:9550630; PMID:7542800
A:Accession: C64076

A:Molecule type: DNA
A:Residues: 1-548 <TIGR>
A:Cross-references: GB:U2736; GB:U42023; NID:g1573519; PIDN:AAC22201.1; PID:g1573528; T

C:Genetics:
A:Gene: groEL
C:Superfamily: chaperonin groEL

C:Keywords: heat shock; molecular chaperone; stress-induced protein

Query Match 68.2%; Score 30; DB 2; Length 548;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 7

Db |||:|
480 NAGTEOY 486

RESULT 36

T43619

secreted protein kinase ypkA - Yersinia pestis plasmid pCD1

N:Alternate names: protein kinase A homolog

C:Species: Yersinia pestis

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T43619; T42850

R:Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnea, J.; Kobayashi, A.; Brubaker

J. Bacteriol. 180, 5192-5202, 1998

A:Reference number: 222578; MUID:96422474; PMID:9748454

A:Accession: T43619

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-732 <HUP>
A:Cross-references: EMBL:AF053946; NID:g2996222; PIDN:AAC62602.1; PID:g2996279

A:Experimental source: strain KIM

R:Berry, R.D.; Straley, S.C.; Fetherston, J.D.; Rose, D.J.; Gregor, J.; Blattner, F.R.

Infect. Immun. 66, 4611-4623, 1998

A:Title: DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersinia p

A:Reference number: 222273; MUID:96427122; PMID:9746557

A:Accession: T42850

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-732 <PER>

A:Cross-references: EMBL:AF074612; NID:g3822037; PIDN:AAC69765.1; PID:g3822045

A:Experimental source: strain KIMS

C:Genetics:

A:Gene: ypkA

A:Genome: plasmid pCD1

Query Match 68.2%; Score 30; DB 2; Length 732;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GTQOYT 8
|||:|
Db 653 GTQOYT 658

RESULT 37

S30060
protein kinase ypkA - Yersinia pseudotuberculosis

C:Species: Yersinia pseudotuberculosis

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999

C:Accession: S30060

R:Galayov, E.E.; Hakansson, S.; Forsberg, A.; Wolf-Watz, H.

Nature 361, 730-733, 1993

A:Title: A secreted protein kinase of Yersinia pseudotuberculosis is an indispensable vi

A:Reference number: S30060; MUID:93180911; PMID:8441468

A:Accession: S30060

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-732 <GAL>
A:Cross-references: EMBL:X69439; NID:g49186; PIDN:CAA49215.1; PID:g49187

C:Keywords: serine/threonine-specific protein kinase

Query Match 68.2%; Score 30; DB 2; Length 732;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GTQOYT 8
|||:|
Db 653 GTQOYT 658

RESULT 38

H95882

probable aldehyde dehydrogenase protein [imported] - Sinorhizobium meliloti (strain 1021

C:Species: *Sinorhizobium meliloti*
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 17-Mar-2003
C:Accession: H95882
R:Finan, T.M.; Weidner, S.; Wong, K.; Bhurmesler, J.; Chai, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: H95882
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-737 <KUR>
A:Cross-references: GB:AL591965; PIDN:CAC48728.1; PID:G15140201; GSPDB:GM00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Gallberg, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
P.; L.; Chai, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federpiet, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Weller, D.H.; Wong, K.; Yeh, K.
Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Comment: annotation
C:Genetics:
A:Gene: SMD20342
A:Genome: Plasmid
C:Superfamily: isoguanidine 1-oxidoreductase beta chain

Query Match 68.2%; Score 30; DB 2; Length 737;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GTCQYT 8
|||:|
Db 227 GTCQYT 232

RESULT 39
T04523
hypothetical protein F16A16.130 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
C:Accession: T04523
R:Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hohenseel, J.; Mew
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15376
A:Accession: T04523
A:Molecule type: DNA
A:Residues: 1-880 <BEV>
A:Cross-references: EMBL:AL035353
A:Experimental source: cultivar Columbia; BAC clone F16A16
C:Genetics:
A:Map position: 4
A:Introns: 660/1
A:Note: F16A16.130

Query Match 68.2%; Score 30; DB 2; Length 880;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTQY 8
|||:|
Db 244 NKGTERT 251

RESULT 40
S30862
DNA dependent ATPase/DNA helicase B - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: Protein YER176w
C:Species: *Saccharomyces cerevisiae*
C:Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 02-Feb-2001
C:Accession: S30862; S50679; J02490; PC2368
R:Nulligan, J.T.; Dietrich, F.S.; Hennessy, K.M.; Schl, P.; Komp, C.; Wei, Y.; Taylor,
submitted to the EMBL Data Library, February 1993

A:Reference number: S30812
A:Accession: S30862
A:Molecule type: DNA
A:Residues: 1-1121 <ML>
A:Cross-references: GB:U18922; EMBL:L11229; NID:G603405; PIDN:AAB64703.1; PID:G603417
R:Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of *S. cerevisiae* cosmid 9163 and 9132.
A:Reference number: S50679
A:Accession: S50679
A:Molecule type: DNA
A:Residues: 1-1121 <DIE>
A:Cross-references: EMBL:U18922; NID:G603405; PIDN:AAB64703.1; PID:G603417; MIPS:YER176w
R:Biswas, E.B.; Chen, P.H.; Leszyk, J.; Biswas, S.B.
Biochem. Biophys. Res. Commun. 206, 850-856, 1995
A:Title: Biochemical and genetic characterization of a replication protein A dependent D
A:Reference number: J02490; MUID:95134267; PMID:7832796
A:Accession: J02490
A:Molecule type: DNA
A:Residues: 1-1121 <BIS>
A:Accession: PC2368
A:Molecule type: protein
A:Residues: 277-283;623-633,'X',635-643 <BI2>
C:Comment: This enzyme plays pivotal roles in the unwinding of the DNA double helix dur
A:Gene: SGD:ECM32
A:Cross-references: SGD:S0000978; MIPS:YER176w
A:Map position: 5R
C:Keywords: nucleotide binding; P-loop
F:670-677/Region: nucleotide-binding motif A (P-loop)

Query Match 68.2%; Score 30; DB 2; Length 1121;
Best Local Similarity 62.5%; Pred. No. 3.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQY 8
|||:|
Db 382 NKGTSRWT 389

Search completed: August 20, 2003, 09:32:40
Job time: 10.9897 secs

DE Trypsin I precursor (EC 3.4.21.4).
 OS Gadus morhua (Atlantic cod).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pyloic caeca;
 RX MEDLINE=94039130; PubMed=8223632;
 RA Gudmundsdottir A., Gudmundsdottir E., Oskarsson S., Bjarnason J.B.,
 RA Rakin A.B., Craik C.S.;
 RT "Isolation and characterization of cDNAs from Atlantic cod encoding
 RT two different forms of trypsinogen.";
 RL Eur. J. Biochem. 217:1091-1097(1993).
 RN [2]
 RP SEQUENCE OF 20-58.
 RC TISSUE=Pyloic caeca;
 MEDLINE=89210867; PubMed=2707266;
 Aagejsson B., Fox J.W., Bjarnason J.B.;
 "Purification and characterization of trypsin from the poikilotherm
 Gadus morhua.";
 RL Eur. J. Biochem. 180:85-94(1989).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg|-Xaa, Lys|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X7686; CAAS4214.1; -.
 DR PIR: S39047; S39047.
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.151; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; Trypsin; 1.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 DR Multigene family.
 FT SIGNAL 1 13
 FT PROPEP 14 19
 FT CHAIN 20 241
 FT ACT_SITE 59 59
 FT ACT_SITE 103 103
 FT ACT_SITE 195 195
 FT DISULFID 26 155
 FT DISULFID 44 60
 FT DISULFID 128 228
 FT DISULFID 135 201
 FT DISULFID 166 180
 FT DISULFID 191 215
 FT SITE 189 189
 FT CONFLICT 25 25
 FT CONFLICT 27 28
 FT CONFLICT 43 43
 FT CONFLICT 49 52
 FT CONFLICT 52 52
 SQ SEQUENCE 241 AA; 25941 MW; 44EC9A0106AD1A68 CRC64;
 Query Match 72.7%; Score 32; DB 1; Length 241;
 Best Local Similarity 71.4%; Pired. No. 12;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 77 NESTROY 83
 RESULT 3
 MREC_BACSU STANDARD; PRT; 290 AA.
 AC 001466;
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Rod shape-determining protein mrec.
 GN MREC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=93015732; PubMed=1400225;
 RA Varley A.W., Stewart G.C.;
 RT "The divIV region of the Bacillus subtilis chromosome encodes
 RT homologs of Escherichia coli septum placement (minCD) and cell shape
 RT (mreBCD) determinants.";
 RL J. Bacteriol. 174:6729-6742(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93015731; PubMed=1400224;
 RA Levin P.A., Margolis P.S., Setlow P., Losick R., Sun D.;
 RT "Identification of Bacillus subtilis genes for septum placement and
 RT shape determination.";
 RL J. Bacteriol. 174:6717-6726(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=9384377; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertoletti M., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Briganti S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conterton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Euteneier K.D., Ewington J., Fabret C., Ferrati E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly B.J., Grandi G.,
 RA Guisepi G., Guy B.P., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holtsappel S., Hosono S., Hujo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Priescan E., Fujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha B., Roche B., Rose M., Sedate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Serró S.J., Serró P., Shin B.S., Solido B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,
 RA Takeuchi M., Yamakoshi A., Tanaka T., Teipstra P., Tognoni A.,
 RA Tosoato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viñari A., Mamout R., Medler B., Medler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: INVOLVED IN FORMATION OF THE ROD SHAPE OF THE CELL. MAY
 CC ALSO CONTRIBUTE TO REGULATION OF FORMATION OF PENICILLIN-BINDING
 CC PROTEINS (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M95582; AAA22606.1; -
 CC EMBL; M96343; AAA22398.1; -
 CC EMBL; Z99118; CAB14762.1; -
 CC PIR; C45240; C45240.
 CC PIR; E69660; E69660.
 CC Subtilist; BG10327; mrec.
 CC InterPro; IPR005223; mrec.
 CC Pfam; PF04085; mrec; 1.
 CC TIGSFAMS; TIGR00219; mrec; 1.
 CC Cell shape; Complete proteome.
 CC CONFLICT 143 143 K -> F (IN REF. 2).
 CC CONFLICT 176 176 T -> P (IN REF. 2).
 CC SEQUENCE 290 AA; 32140 MW; B9984D25B8414A5C CRC64;

Very Match 72.7%; Score 32; DB 1; Length 290;
 Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOQ 6
 134 NKGTOQ 139

RESULT 4

PHYD ARATH STANDARD; PRT; 1164 AA.

AC P42497; Q23472; 32. Created)
 DT 01-NOV-1995 (Rel. 32. Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phycochrome D.
 GN PHYD OR AT4G16250 OR DL4165C.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eusteroideae; Brassicales; Brassicaceae; Arabidopsids.
 CC NCBI TaxID=3702;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RX MEDLINE=94325466; PubMed=8049367;
 P "Clack T., Matthews S., Sharrock R.A.;
 "The phytochrome apoprotein family in Arabidopsis is encoded by five
 genes: the sequences and expression of PHYD and PHYE.";
 RL Plant Mol. Biol. 25:413-427(1994).
 [2]
 SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=9812113; PubMed=9461215;
 RA Bevan M., Bancroft I., Bent E., Love K., Goodman H.M., Dean C.,
 Ridley P., Hudson S.-A., Patel K., Murphy G., Piffenelli P.,
 Wedler H., Medler E., Wambut R., Weitzenecker T., Pohl T., Terryn N.,
 Gielen J., Villarroel R., De Clercq R., van Montagu M., Lecharny A.,
 Aubourg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S.,
 Kötter P., Entian K.-D., Rieger M., Schaefer M., Funk B.,
 Muller-Auer S., Silvey M., James R., Monfort A., Pons A.,
 Puzosmeh P., Douka A., Voukeltou E., Milioni D., Hatzopoulos P.,
 Pivaardi E., Obermaier B., Hilbert H., Duesterhoef A., Moeres T.,
 Jones J.D.G., Eneva T., Palme K., Benes V., Rechmann S., Ansoerge W.,
 Cooke R., Berger C., Delany M., Voet M., Volckaert G., Mewes H.-W.,
 Klosteman S., Schueller C., Chalmers N.,
 RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
 Arabidopsis thaliana";
 RL Nature 391:485-488(1998).
 [3]
 SEQUENCE FROM N.A.

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

RC STRAIN=cv. Columbia;

RX MEDLINE=2008348; PubMed=10617198;

RA Mayer K.F.X., Schueller C., Wambut R., Murphy G., Volckaert G.,

RA Pohl T., Duesterhoef A., Stiekema W., Entian K.-D., Terryn N.,

RA Harris B., Ansoerge W., Brandt P., Grivell L.A., Rieger M.,

RA Weissegger T., de Simone V., Obermaier B., Meche R., Mueller M.,

RA Kreis M., Delany M., Puzosmeh P., Watson M., Schidheini T.,

RA Reichert B., Portetle D., Perez-Alonso M., Boutry M., Bancroft I.,

RA Vos P., Hobeisel J., Zimmermann W., Medler H., Ridley P.,

RA Langham S.-A., McCullagh B., Bilham L., Robben J.,

RA Van der Schueren J., Grymoprez B., Chuang Y.-D., Vandenbusche F.,

RA Braeken M., Weitzenecker T., Bothe G., Rameberger U., Hilbert H., Braun M.,

RA Holzer E., Brandt A., Peters S., van Staveren M., Dirke P.,

RA Moijman P., Klein Lankhorst R., Rose M., Hauf J., Kötter P.,

RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,

RA De Keyser A., Buyschaert C., Gielen J., Quail M., Bray-Allen S.,

RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,

RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,

RA Petter A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,

RA Borova D., Bloeker H., Scharfe M., Grimm M., Loebner T.-H.,

RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,

RA Neumann S., Argitlou A., Vitale D., Lignori R., Pivaardi E.,

RA Massenet O., Quigley F., Clabaud G., Mendlein A., Falber R.,

RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,

RA Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,

RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,

RA Perez-Perez A., Furnelle B., Bent E., Johnson S., Tacon D., Jesse T.,

RA Heljens L., Schwarz S., Scholler P., Heber S., Francis P., Bleike C.,

RA Frieleman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,

RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,

RA Parnell L., Dedbia N., Gnoj L., Schütz K., Huang E., Spiegel L.,

RA Sektion M., Murray J., Shee P., Cordes M., Abu-Threidat J.,

RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,

RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,

RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,

RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,

RA Nelson J., Speith J., Ryan E., Andrews S., Geisel C., Layman D.,

RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,

RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,

RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,

RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,

RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,

RA Chen E., Marra M., Martensen R., McConbie W.R.,

RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis

thaliana";

RL Nature 402:769-777(1999).

CC -! FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT

ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS

MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT

ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN

PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS

RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE

RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR

GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-

BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,

PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS

THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.

-! SUBUNIT: Homodimer.

-! PTM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.

-! SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.

-! SIMILARITY: Contains 1 histidine kinase domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X76609; CAA54072.1; -
 DR EMBL; Z97340; CAB10404.1; -
 DR PIR; A161543; CAB78667.1; -
 DR PIR; B71429; B71429.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR003018; GAF.
 DR InterPro; IPR003661; His_kinase.
 DR InterPro; IPR005467; His_kinase.
 DR InterPro; IPR000014; PAS_domain.
 DR InterPro; IPR001294; Phytochrome.
 DR Pfam; PF01590; GAF_1.
 DR Pfam; PF02518; HATPase_C_1.
 DR Pfam; PF00512; HisKA_1.
 DR Pfam; PF00369; PAS_2.
 DR Pfam; PF00360; Phytochrome_1.
 DR PRINTS; PR01033; PHYTOCHROME.
 DR SMART; SM00065; GAF_1.
 DR SMART; SM00387; HATPase_C_1.
 DR SMART; SM00388; HisKA_1.
 DR SMART; SM00091; PAS_2.
 DR TIGRfam; TIGR00229; sensory_box_2.
 DR PROSITE; PS50109; HIS_KIN_1.
 DR PROSITE; PS50112; PAS_2.
 DR PROSITE; PS00245; PHYTOCHROME_1_1.
 DR PROSITE; PS50046; PHYTOCHROME_2_1.
 DR Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
 KW Repeat; Multigene family.
 FT DOMAIN 656 727 PAS 1.
 FT DOMAIN 790 861 PAS 2.
 FT BINDING 938 1157 HISTIDINE KINASE.
 FT BINDING 360 360 CHROMOPHORE (BY SIMILARITY).
 FT CONFLICT 425 425 L -> F (IN REF. 1).
 SQ SEQUENCE 1164 AA; 129267 MW; BB7CBE19C50ACB4B CRC64;

Query Match 72.7%; Score 32; DB 1; Length 1164;
 Best Local Similarity 75.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
 DB 57 NKAIQOYT 64

RESULT 5
 YPOL ACICA STANDARD; PRT; 204 AA.
 AC P07778; 01-AUG-1988 (Rel. 08, Created)
 01-AUG-1988 (Rel. 08, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein in PQO-V 5' region (ORF L) (Fragment).
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LMD 79.41;
 RX MEDLINE=89123056; PubMed=2536663;
 RA Goosen N., Hoxman H.P.A., Huinen R.G.M., van de Putte P.;
 RT "Acinetobacter calcoaceticus genes involved in biosynthesis of the
 coenzyme pyrolo-quinoline-quinone: nucleotide sequence and
 expression in *Escherichia coli* K-12.";
 RL J. Bacteriol. 171:447-455 (1989).
 CC -1- FUNCTION: NOT KNOWN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X06452; CAA29752.1; -
 DR PIR; A32252; A32252.
 DR InterPro; IPR004843; M-peptidase.
 DR Pfam; PF00149; Metallophos; 1.
 KW Hypothetical protein.
 FT NON TER 204 204
 SQ SEQUENCE 204 AA; 22132 MW; AB5E0D4E7988AB84 CRC64;

Query Match 70.5%; Score 31; DB 1; Length 204;
 Best Local Similarity 75.0%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
 DB 152 NKGCQOYT 159

RESULT 6
 SSRA ONCMY STANDARD; PRT; 288 AA.
 AC P45433; 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Translocon-associated protein, alpha subunit precursor (TRAP-alpha)
 DE (Signal sequence receptor alpha subunit) (SSR-alpha).
 GN SSRI.
 OS Oncorhynchus mykiss (Rainbow trout) (*Salmo gairdneri*).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Procaracanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94326944; PubMed=8050590;
 RA Hartmann E., Prehn S.;
 RT "The N-terminal region of the alpha-subunit of the TRAP complex has a
 conserved cluster of negative charges.";
 RL FEBS Lett. 349:324-326 (1994).
 CC -1- FUNCTION: TRAP PROTEINS ARE PART OF A COMPLEX WHOSE FUNCTION IS TO
 CC BIND CA(2+) TO THE ER MEMBRANE AND THEREBY REGULATE THE RETENTION
 CC OF ER RESIDENT PROTEINS. MAY BE INVOLVED IN THE RECYCLING OF THE
 CC TRANSLOCATION APPARATUS AFTER COMPLETION OF THE TRANSLOCATION
 CC PROCESS OR MAY FUNCTION AS A MEMBRANE-BOUND CHAPERONE FACILITATING
 CC FOLDING OF TRANSLOCATED PROTEINS.
 CC -1- SUBUNIT: HETEROTETRAMER OF TRAP-ALPHA, TRAP-BETA, TRAP-DELTA AND
 CC TRAP-GAMMA.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
 CC reticulum.
 CC -1- DOMAIN: SHOWS A REMARKABLE CHARGE DISTRIBUTION WITH THE N-TERMINUS
 CC BEING HIGHLY NEGATIVELY CHARGED, AND THE CYTOPLASMIC C-TERMINUS
 CC POSITIVELY CHARGED.
 CC -1- PTM: PHOSPHORYLATED IN ITS CYTOPLASMIC TAIL (BY SIMILARITY).
 CC -1- MISCELLANEOUS: SEEMS TO BIND CALCIUM.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; Z12831; CAA78291.1; -
 DR PIR; I51332; I51332.
 DR InterPro; IPR005595; TRAP_alpha.
 DR Pfam; PF03896; TRAP_alpha_1.
 KW Glycoprotein; Signal; Phosphorylation; Endoplasmic reticulum;
 KW Transmembrane; Calcium-binding.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 288 TRANSLOCON-ASSOCIATED PROTEIN, ALPHA

FT DOMAIN 34 207 SUBUNIT.
 FT TRANSMEM 208 228 LUMENAL (POTENTIAL).
 FT DOMAIN 229 288 POTENTIAL.
 FT CARBOHYD 137 137 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 132 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 288 AA; 31855 MW; 84723DBF17FD7C8E CRC64;
 Query Match 70.5%; Score 31; DB 1; Length 288;
 Best Local Similarity 62.5%; Pred. No. 23;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTOQYT 8
 DB 110 NKGQDQFT 117
 RESULT 7
 P11A MOUSE STANDARD; PRT; 1068 AA.
 P42337;
 01-NOV-1995 (Rel. 32, Last sequence update)
 01-NOV-1995 (Rel. 32, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit,
 DE alpha isoform (EC 2.7.1.153) (PI3-kinase p110 subunit alpha) (Ptdins-
 DE 3-kinase p110) (PI3K).
 GN PI3KA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/C;
 RA MEDLINE=94187738; PubMed=8139567;
 RA Klippel A., Escobedo J.A., Hirano M., Williams L.T.;
 RT "The interaction of small domains between the subunits of
 RT phosphatidylinositol 3-kinase determines enzyme activity.";
 RL Mol. Cell Biol. 14:2675-2685(1994).
 CC -1- FUNCTION: PHOSPHORYLATES PTDINS, PTDINS4P AND PTDINS(4,5)P2 WITH A
 CC PREFERENCE FOR PTDINS(4,5)P2.
 CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4,5-
 CC bisphosphate = ADP + 1-phosphatidyl-1D-myo-inositol 3,4,5-
 CC triphosphate.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U03279; AAA18334.1; -
 DR MGD; MGI:1206581; PI3Ka.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR000403; PI3_P14_kinase.
 DR InterPro: IPR002420; PI3K_C2_
 DR InterPro: IPR003113; PI3K_p85B.
 DR InterPro: IPR00341; PI3K_ras_bind.
 DR InterPro: IPR001263; PI3Ka.
 DR Pfam; PF00454; P13_P14_kinase; 1.
 DR Pfam; PF00792; PI3K_C2; 1.
 DR Pfam; PF02192; PI3K_p85B; 1.
 DR Pfam; PF00794; PI3K_rbd; 1.
 DR Pfam; PF00613; PI3Ka; 1.
 DR SMART; SM00232; C2; 1.
 DR SMART; SM00142; PI3K_C2; 1.

DR SMART; SM00143; PI3K_p85B; 1.
 DR SMART; SM00144; PI3K_rbd; 1.
 DR SMART; SM00145; PI3Ka; 1.
 DR SMART; SM00146; PI3K; 1.
 DR PROSITE; PS50004; C2_DOMAIN_2; FALSE_NEG.
 DR PROSITE; PS00915; P13_4_KINASE_1; 1.
 DR PROSITE; PS00916; P13_4_KINASE_2; 1.
 DR PROSITE; PS0290; P13_4_KINASE_3; 1.
 KW Transferase; Kinase; Multigene family.
 FT DOMAIN 319 428 C2 DOMAIN.
 FT DOMAIN 757 1068 PI3K/P14K.
 SQ SEQUENCE 1068 AA; 124453 MW; 67F54FCF8747FB1C CRC64;
 Query Match 70.5%; Score 31; DB 1; Length 1068;
 Best Local Similarity 62.5%; Pred. No. 90;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTOQYT 8
 DB 965 SKQAQBYT 972
 RESULT 8
 BRR2 YEAST STANDARD; PRT; 2163 AA.
 ID BRR2 YEAST
 AC P32639;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Pre-mRNA splicing helicase BRR2 (EC 3.6.1.-) (Protein Snu246).
 GN BRR2 OR RSS1 OR SNU246 OR YER172C OR SYCP-ORF66.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 RA Arayo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Guzman T., Hartzell G.,
 RA Huntke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oeffner P.,
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
 RL Nature 387:78-81(1997).
 RN [2]
 RP CHARACTERIZATION.
 RC STRAIN=DRY473;
 RX MEDLINE=96304576; PubMed=8722763;
 RA Noble S.M., Guthrie C.;
 RT "Identification of novel genes required for yeast pre-mRNA splicing
 RT by means of cold-sensitive mutations.";
 RL Genetics 143:67-80(1996).
 RN [3]
 RP CHARACTERIZATION.
 RC MEDLINE=96324408; PubMed=8670905;
 RA Lauber J., Fabrizio P., Teigelkamp S., Lane W.S., Hartmann E.,
 RA Luehrmann R.;
 RT "The HeLa 200 kDa US snRNP-specific protein and its homologue in
 RT Saccharomyces cerevisiae are members of the DEXH-box protein family of
 RT putative RNA helicases.";
 RL EMBO J. 15:4001-4015(1996).
 CC -1- FUNCTION: REQUIRED FOR PRE-MRNA SPLICING.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- DOMAIN: COMPOSED OF TWO SIMILAR DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. SKI2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U18922; AAB64699.1; -.

DR PIR; S50675; S50675.

DR SGD; S0000974; BR2.

DR GO; GO:0005687; C:snrnp U4; IDA.

DR GO; GO:0005682; C:snrnp U5; IDA.

DR GO; GO:0005688; C:snrnp U6; IDA.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR004179; Sec63.

DR Pfam; PF00270; DEAD_2.

DR Pfam; PF00271; helicase_C.1.

DR Pfam; PF02889; Sec63_2.

DR SMART; SM00382; AAA; 2.

DR SMART; SM00487; DEXDC; 2.

DR SMART; SM00490; HELIC; 1.

DR SMART; SM00611; SEC63; 2.

DR Hydroxylase; Helicase; mRNA processing; mRNA splicing; Spliceosome;

DR Nucleic acid; ATP-binding; Repeat.

FT DOMAIN 521 923

FT NP_BIND 1370 1757

FT NP_BIND 521 528

FT NP_BIND 1370 1377

FT SITE 634 637

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

Query Match 68.2%; Score 30; DB 1; Length 60;
Best Local Similarity 62.5%; Pred. No. 7.2;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

CC -----

DR EMBL; U18922; AAB64699.1; -.

DR PIR; S50675; S50675.

DR SGD; S0000974; BR2.

DR GO; GO:0005687; C:snrnp U4; IDA.

DR GO; GO:0005682; C:snrnp U5; IDA.

DR GO; GO:0005688; C:snrnp U6; IDA.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR004179; Sec63.

DR Pfam; PF00270; DEAD_2.

DR Pfam; PF00271; helicase_C.1.

DR Pfam; PF02889; Sec63_2.

DR SMART; SM00382; AAA; 2.

DR SMART; SM00487; DEXDC; 2.

DR SMART; SM00490; HELIC; 1.

DR SMART; SM00611; SEC63; 2.

DR Hydroxylase; Helicase; mRNA processing; mRNA splicing; Spliceosome;

DR Nucleic acid; ATP-binding; Repeat.

FT DOMAIN 521 923

FT NP_BIND 1370 1757

FT NP_BIND 521 528

FT NP_BIND 1370 1377

FT SITE 634 637

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

FT SITE 1474 1477

Query Match 68.2%; Score 30; DB 1; Length 258;
 Best Local Similarity 71.4%; Pred. No. 33;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
 DB 83 NDGTQY 89

RESULT 11

ELI_BOVIN
 ID ELI_BOVIN STANDARD; PRT; 266 AA.
 AC Q28153;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Elastase 1 precursor (EC 3.4.21.36).
 GN ELA1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;

OX NCB1_TaxID=9913;
 RN SEQUENCE FROM N.A.
 RC STRAIN=Holstein-Friesian; TISSUE=Pancreas;
 RX MEDLINE=98079203; PubMed=9418008;
 RA Gestin M., le Huecrou-Luron I., Wicker-Planquart C., le Drian G.,
 Chaix J.C., Puliguer A., Guilloteau P.,
 RT "Bovine pancreatic preproelastases I and II: comparison of nucleotide
 and amino acid sequences and tissue specific expression."
 RL Comp. Biochem. Physiol. 118B:181-187(1997).
 CC -1- FUNCTION: ACTS UPON ELASTIN.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins, including elastin.
 CC -1- Preferential cleavage: Ala-Xaa.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Pancreas.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----

CC EMBL; M80838; AAA98525.1; -.
 CC HSSP; P00772; 1ELG.
 CC MEROPS; S01.153; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SMO0020; TRYSPC; 1.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PSS0134; TRYPSIN_HIS; 1.
 DR PROSITE; PSS0135; TRYPSIN_SER; 1.
 KM Hydrolyase; Serine protease; Zymogen; Signal.
 FT SIGNAL 1 16
 FT PROPEP 17 26
 FT CHAIN 27 266
 FT ACT_SITE 71 71
 FT ACT_SITE 119 119
 FT ACT_SITE 214 214
 FT DISULFID 56 72
 FT DISULFID 153 220
 FT DISULFID 184 200
 FT DISULFID 210 240
 FT CARBOHYD 87 87
 FT CARBOHYD 241 241
 FT SEQUENCE 266 AA; 28518 MW; 927E29C69BAF67E7 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 266;
 Best Local Similarity 71.4%; Pred. No. 35;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
 DB 91 NDGTQY 97

RESULT 12

ELI_PIG
 ID ELI_PIG STANDARD; PRT; 266 AA.
 AC P00772; Q29625;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Elastase 1 precursor (EC 3.4.21.36).
 GN ELA1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 NCBI_TaxID=9823;

OX NCB1_TaxID=9823;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=86304235; PubMed=3528137;
 RA Shirasu Y., Yoshida H., Mikayama T., Matsuki S., Tanaka J.,
 RT "Isolation and expression in Escherichia coli of a cDNA clone
 encoding porcine pancreatic elastase."
 RL J. Biochem. 99:1707-1712(1986).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=87250343; PubMed=3648024;
 RA Tani T., Kawahara I., Furukawa H., Ohmine T., Takiguchi Y.,
 RT "Characterization of a silent gene for human pancreatic elastase I:
 structure of the 5'-flanking region."
 RL J. Biochem. 101:591-599(1987).
 RN [3]
 RN SEQUENCE OF 27-266.
 RP MEDLINE=73229121; PubMed=4578945;
 RA Shotton D.M., Hartley B.S.,
 RT "Evidence for the amino acid sequence of porcine pancreatic
 elastase."
 RL Biochem. J. 131:643-675(1973).
 RN [4]
 RN SEQUENCE OF 27-266.
 RP MEDLINE=70114042; PubMed=5415108;
 RA Shotton D.M., Hartley B.S.,
 RT "Amino-acid sequence of porcine pancreatic elastase and its
 homologues with other serine proteinases."
 RL Nature 225:802-806(1970).
 RN [5]
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RP MEDLINE=78111463; PubMed=628010;
 RA Sawyer L., Shotton D.M., Campbell J.W., Wendell P.L., Muirhead H.,
 RA Watson H.C., Diamond R., Ladner R.C.,
 RT "The atomic structure of crystalline porcine pancreatic elastase at
 2.5-A resolution: comparisons with the structure of
 alpha-chymotrypsin."
 RL J. Mol. Biol. 118:137-208(1978).
 RN [6]
 RN X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS), AND ACTIVE SITE.
 RP MEDLINE=70114044; PubMed=5415110;
 RA Shotton D.M., Watson H.C.,
 RT "Three-dimensional structure of tosyl-elastase."
 RL Nature 225:811-816(1970).
 RN [7]
 RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF COMPLEX WITH INHIBITOR.
 RP MEDLINE=95006335; PubMed=7922044;
 RA Huang K., Strynadka N.C., Bernard V.D., Peanasky R.J., James M.N.,
 RT "The molecular structure of the complex of Ascaris
 chymotrypsin/elastase inhibitor with porcine elastase."
 RL Structure 2:679-689(1994).

```

CC -1- FUNCTION: ACTS UPON ELASTIN.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins, including elastin.
CC -1- PREFERENTIAL CLEAVAGE: Ala-I-Xaa.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Pancreas.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -1- DATABASE: NMBE-WORTHINGTON ENZYME MANUAL;
CC WWW-http://www.worthington-biochem.com/manual/E/ES.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL; X04036; CAA27670.1; -.
CC EMBL; D00160; BAA00118.1; -.
CC PIR; JS0013; ELPG.
CC PDB; 1EST; 16-OCT-87.
CC PDB; 2EST; 30-APR-94.
CC PDB; 3EST; 15-OCT-92.
CC PDB; 4EST; 15-JUL-92.
CC PDB; 5EST; 15-APR-92.
CC PDB; 6EST; 15-OCT-91.
CC PDB; 7EST; 30-APR-94.
CC PDB; 8EST; 15-JAN-93.
CC PDB; 9EST; 31-JAN-94.
CC PDB; 1INC; 31-JAN-94.
CC PDB; 1JTM; 31-JAN-94.
CC PDB; 1ELA; 30-APR-94.
CC PDB; 1ELB; 22-JUN-94.
CC PDB; 1ELC; 30-APR-94.
CC PDB; 1ELD; 14-FEB-95.
CC PDB; 1ELE; 14-FEB-95.
CC PDB; 1ELF; 10-JUL-95.
CC PDB; 1ELG; 10-JUL-95.
CC PDB; 1ESA; 30-APR-94.
CC PDB; 1ESB; 30-APR-94.
CC PDB; 1EAT; 05-APR-99.
CC PDB; 1EAS; 07-FEB-95.
CC PDB; 1EAT; 07-FEB-95.
CC PDB; 1EAV; 07-FEB-95.
CC PDB; 1BMA; 07-DEC-95.
CC PDB; 1LVY; 27-JAN-97.
CC PDB; 1NES; 29-JAN-96.
CC PDB; 1BOE; 18-NOV-98.
CC PDB; 1BTU; 16-FEB-99.
CC PDB; 1CJM; 28-JUL-99.
CC PDB; 1B34; 18-OCT-00.
CC PDB; 1B35; 18-OCT-00.
CC PDB; 1B36; 18-OCT-00.
CC PDB; 1B37; 18-OCT-00.
CC PDB; 1B38; 18-OCT-00.
CC PDB; 1PLE; 11-JAN-97.
CC PDB; 1PZZ; 04-OCT-01.
CC PDB; 1GWA; 06-JUN-02.
CC PDB; 1HAZ; 02-AUG-01.
CC PDB; 1HV7; 02-MAY-01.
CC PDB; 1LOZ; 28-AUG-02.
CC PDB; 1LIG; 28-AUG-02.
CC PDB; 1LKA; 28-AUG-02.
CC PDB; 1LKB; 28-AUG-02.
CC PDB; 1MCV; 04-FEB-03.
CC PDB; 1MMJ; 23-DEC-02.
CC PDB; 1QGF; 06-DEC-00.
CC PDB; 1QIX; 14-DEC-99.
CC PDB; 1ONJ; 25-APR-00.
CC PDB; 1QR3; 09-OCT-02.
CC MEROPS; S01.153; -.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.

```

```

DR SMART; SM00020; TRYP_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Zymogen; Signal; 3D-structure.
FT SIGNAL 1 16
FT PROPEP 17 26
FT CHAIN 27 266
FT ACT_SITE 71 71
FT ACT_SITE 119 119
FT ACT_SITE 214 214
FT DISULFID 56 72
FT DISULFID 153 220
FT DISULFID 184 200
FT DISULFID 210 240
FT CONFLICT 92 92
FT CONFLICT 126 126
FT CONFLICT 184 184
FT CONFLICT 204 204
FT STRAND 28 28
FT TURN 29 29
FT STRAND 31 32
FT TURN 35 36
FT TURN 39 40
FT STRAND 41 48
FT TURN 49 50
FT STRAND 51 62
FT TURN 63 64
FT STRAND 65 68
FT HELIX 70 73
FT STRAND 79 83
FT STRAND 87 87
FT TURN 88 89
FT STRAND 96 105
FT TURN 107 108
FT TURN 111 112
FT HELIX 114 116
FT STRAND 121 125
FT STRAND 132 132
FT TURN 133 134
FT STRAND 135 135
FT TURN 143 144
FT TURN 149 150
FT STRAND 153 157
FT STRAND 160 160
FT TURN 163 164
FT STRAND 167 167
FT STRAND 170 170
FT STRAND 172 175
FT STRAND 178 179
FT HELIX 181 184
FT TURN 185 185
FT TURN 187 190
FT HELIX 191 193
FT TURN 196 197
FT STRAND 198 198
FT STRAND 201 201
FT STRAND 208 208
FT TURN 211 212
FT TURN 214 215
FT STRAND 217 222
FT TURN 223 224
FT STRAND 225 234
FT STRAND 236 236
FT TURN 237 238
FT STRAND 239 239

```

Query Match 68.2%; Score 30; DB 1; Length 266;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
Db 91 NDGTQOY 97

RESULT 13

ELI_RAT	STANDARD	PRT	266 AA
AC	P00773		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Elastase 1 precursor (EC 3.4.21.36).		
GN	ELAI.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxId=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=82182967; PubMed=6918221;		
RA	McDonald R.J., Swift G.H., Quinto C., Swain W., Pictet R.L.,		
RA	Nikolovits W., Rutter W.J.;		
RT	"Primary structure of two distinct rat pancreatic preproelastases determined by sequence analysis of the complete cloned messenger ribonucleic acid sequences."		
RT	Biochemistry 21:1453-1463(1982).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=85054882; PubMed=6094548;		
RA	Swift G.H., Craik C.S., Stary S.J., Quinto C., Lahaie R.G.,		
RA	Rutter W.J., McDonald R.J.;		
RT	"Structure of the two related elastase genes expressed in the rat pancreas."		
RT	J. Biol. Chem. 259:14271-14278(1984).		
RN	[3]		
RP	SEQUENCE OF 17-45.		
RC	TISUS=Pancreeas;		
RX	MEDLINE=84000385; PubMed=6555050;		
RA	Largman C.;		
RT	"Isolation and characterization of rat pancreatic elastase."		
RT	Biochemistry 22:3763-3770(1983).		
CC	-1- FUNCTION: ACTS UPON ELASTIN.		
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of proteins, including elastin.		
CC	-1- PREFERENTIAL CLEAVAGE: Ala-Xaa.		
CC	-1- SUBCELLULAR LOCATION: Secreted.		
CC	-1- TISSUE SPECIFICITY: Pancreas.		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; VO1234; CAA24544.1; -		
DR	EMBL; L00117; AAA98811.1; -		
DR	EMBL; L00112; AAA98811.1; JOINED.		
DR	EMBL; L00113; AAA98811.1; JOINED.		
DR	EMBL; L00114; AAA98811.1; JOINED.		
DR	EMBL; L00115; AAA98811.1; JOINED.		
DR	EMBL; L00116; AAA98811.1; JOINED.		
DR	PIR; A00960; ELRT1.		
DR	HSSP; P00772; IELG.		
DR	MEROPS; S01.153; -		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	Pfam; PF00089; trypsin.1.		
DR	SMART; SM00020; Tryp_SPC; 1.		
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.		
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.		
DR	PROSITE; PS00135; TRYPSIN_SER; 1.		
KW	Hydrolase; Serine protease; Zymogen; Signal.		
FT	SIGNAL 1 16		
FT	PROPEP 17 26 ACTIVATION PEPTIDE.		

FT CHAIN 27 266 ELASTASE 1.

FT ACT_SITE 71 71 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 119 119 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 56 72 BY SIMILARITY.

FT DISULFID 153 220 BY SIMILARITY.

FT DISULFID 184 200 BY SIMILARITY.

FT DISULFID 210 240 BY SIMILARITY.

FT CONFLICT 104 104 M -> V (IN REF. 2).

FT CONFLICT 108 108 T -> N (IN REF. 2).

FT CONFLICT 244 244 K -> R (IN REF. 2).

FT CONFLICT 266 266 T -> N (IN REF. 2).

SQ SEQUENCE 266 AA; 28976 MW; 5A56FE8FCF1AEDA CRC64;

Query Match

Best Local Similarity 68.2%; Score 30; DB 1; Length 266;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTOQY 7

Db 91 NDGTEQY 97

RESULT 14

IGC2_HUMAN	STANDARD	PRT	326 AA
ID	P01859;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Ig gamma-2 chain C region.		
GN	IGHG2.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE OF 2-326 FROM N.A.		
RX	MEDLINE=82197621; PubMed=6804948;		
RA	Ellison J.W., Hood L.E.;		
RT	"Linkage and sequence homology of two human immunoglobulin gamma heavy chain constant region genes."		
RT	Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).		
RN	[2]		
RP	SEQUENCE OF 88-115 FROM N.A.		
RC	TISUS=Fetal liver;		
RX	MEDLINE=83001943; PubMed=6811139;		
RA	Takahashi N., Ueda S., Obata M., Nikaido T., Nakai S., Honjo T.;		
RT	"Structure of human immunoglobulin gamma genes: implications for evolution of a gene family."		
RT	Cell 29:671-679(1982).		
RN	[3]		
RP	SEQUENCE OF 99-177 AND 310-326 FROM N.A.		
RC	TISUS=Fetal liver;		
RX	MEDLINE=84235992; PubMed=6329676;		
RA	Krawinkel U., Rabbits T.H.;		
RT	"Comparison of the hinge-coding segments in human immunoglobulin gamma heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass genes."		
RT	EMBO J. 1:403-407(1982).		
RN	[4]		
RP	SEQUENCE OF 1-325 (MYELOMA PROTEIN TIT).		
RX	MEDLINE=81007873; PubMed=6774012;		
RA	Wang A.-C., Tung E., Fudenberg H.H.;		
RT	"The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and functional implications."		
RT	J. Immunol. 125:1048-1054(1980).		
RN	[5]		
RP	SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).		
RX	MEDLINE=80001357; PubMed=113060;		
RA	Connell G.E., Parr D.M., Hofmann T.;		
RT	"The amino acid sequences of the three heavy chain constant region domains of a human IgG2 myeloma protein."		

Can. J. Biochem. 57:758-767(1979).
 [6] SEQUENCE OF 238-275 (ZIE).
 RX MEDLINE=80114419; PubMed=118920;
 RA Hofmann T., Part D.M.;
 RT "A note of the amino acid sequence of residues 381-391 of human
 immunoglobulin gamma chains.";
 RL Mol. Immunol. 16:923-925(1979).
 (17)
 RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).
 RA Hofmann T., Part D.M.;
 RL Submitted (MAR-1980) to the PIR data bank.
 (8)
 RP SEQUENCE OF 1-121 (DOT).
 RX MEDLINE=9525298; PubMed=7737190;
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver P., Ferri G.;
 RT "Characterization of the two unique human anti-flavin monoclonal
 immunoglobulins.";
 RL Eur. J. Biochem. 228:886-893(1995).
 (9)
 CC DISULFIDE BONDS.
 RX MEDLINE=72033500; PubMed=4940472;
 RA Miletic C., Frangione B.;
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";
 RL Biochem. J. 121:217-225(1971).
 (10)
 RP DISULFIDE BONDS.
 RX MEDLINE=69064124; PubMed=5782707;
 RA Frangione B., Miletic C., Pink J.R.L.;
 RT "Structural studies of immunoglobulin G.";
 RL Nature 221:145-148(1969).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, J00230; AAB59393.1; --
 DR PIR; A93906; G2HU.
 DR HSSP; P01857; 1FC1.
 DR Genew; HGNC:5526; IGHG2.
 DR MIM; 147110; --
 DR GO; GO:0005624; C:membrane fraction; NAS.
 DR GO; GO:0003823; F:antigen binding activity; TAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-1like.
 DR InterPro; IPR003597; Ig C1.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig_3.
 DR SMART; SM00407; IgC1; 2.
 DR PROSITE; PS50835; IG_LIKE; 3.
 DR PROSITE; PS00290; IG_MHC; 2.
 DR Immunoglobulin domain; Immunoglobulin C region.
 FT NON TER 1 1
 FT DOMAIN 1 96
 FT DOMAIN 99 110
 FT DOMAIN 111 219
 FT DOMAIN 220 326
 FT DISULFID 14 14
 FT DISULFID 27 83
 FT DISULFID 102 102
 FT DISULFID 103 103
 FT DISULFID 106 106
 FT DISULFID 109 109
 FT DISULFID 140 200
 FT DISULFID 246 304
 FT SITE 156 156
 FT MOD_RES 326 326
 FT VARIANT 60 60
 AT OR NEAR THE COMPLEMENT-BINDING SITE.
 REMOVED POST-TRANSLATIONALLY (PROBABLE).
 S -> A (IN MYELOMA PROTEINS TIL & ZIE).
 /FTID=VAR_003889.

FT CONFLICT 109 109 C -> S (IN REF. 3).
 SQ SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;
 Query Match 68.2%; Score 30; DB 1; Length 326;
 Best Local Similarity 75.0%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NKGTQOYT 8
 | | | | |
 DB 75 NKGTQYTT 82

RESULT 15

PR12_MOUSE STANDARD; PRT; 505 AA.

AC P33610;
 DT 01-FEB-1994 (rel. 28; Created)
 DT 01-FEB-1994 (rel. 28; Last sequence update)
 DT 15-JUL-1998 (rel. 36; Last annotation update)
 DE DNA primase large subunit (EC 2.7.7.-) (DNA primase 58 kDa subunit)
 DE (p56).
 GN PRIM2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_Taxid=10090;
 (1)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-48; 186-212 AND 403-416.
 RX MEDLINE=93216788; PubMed=8463324;
 RA Miyazawa H., Izumi M., Tada S., Takada R., Masutani M., Ui M.,
 RA Hanoka F.;
 RT "Molecular cloning of the cDNA for the four subunits of mouse DNA
 polymerase alpha-primase complex and their gene expression during
 cell proliferation and the cell cycle.";
 RL J. Biol. Chem. 268:8111-8122(1993).
 (2)
 RP SEQUENCE FROM N.A.
 RA Bruchner A., Stadlbauer F., Teeng B.Y., Naehue H.P.;
 RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
 RNA PRIMERS FOR THE OKAZAKI FRAGMENTS MADE DURING DISCONTINUOUS
 DNA REPLICATION.
 CC -1- SUBUNIT: HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC PRIMASE LARGE SUBUNIT
 FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; D13545; BAA02745.1; --
 DR EMBL; D17385; BAA04203.1; --
 DR PIR; C46642; C46642.
 DR MGD; MG1:97758; PRIM2.
 DR Pfam; PF04104; DNA_Primase_1; 1.
 DR Transfaser; DNA replication; DNA-directed RNA polymerase; Primosome;
 KW DNA-binding.
 FT CONFLICT 109 109 A -> G (IN REF. 2).
 SQ SEQUENCE 505 AA; 58408 MW; 36A64C77669D8C24 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 505;
 Best Local Similarity 83.3%; Pred. No. 68;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KGTQOY 7
 | | | | |
 DB 65 KGTQOY 70

RESULT 16
 ID PR12_HUMAN STANDARD; PRT; 509 AA.
 AC P49643;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA primase large subunit (EC 2.7.7.-) (DNA primase 58 kDa subunit) (p58).
 DE PR12A OR PR12.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94298818; PubMed=8026492;
 RA Stadlbauer F., Brueckner A., Reithues C., Eckerskorn C.,
 RA Lottspeich F., Foerster V., Tseng B.Y., Nasheuer H.P.;
 DE "DNA replication in vitro by recombinant DNA-polymerase-alpha-primase.";
 CC Eur. J. Biochem. 222:781-793 (1994).
 CC -1- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL RNA PRIMERS FOR THE OKAZAKI FRAGMENTS MADE DURING DISCONTINUOUS DNA REPLICATION.
 CC -1- SUBUNIT: HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC PRIMASE LARGE SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X74331; CAA52378.1; -.
 DR PIR: S45631; S45631.
 DR GeneW: HGNC:9370; PRIM2A.
 DR GK: P49643; -.
 DR MIM: 176636; -.
 DR GO: GO:0003896; F:DNA primase activity; TAS.
 DR GO: GO:0006269; P:DNA replication, priming; TAS.
 DR Pfam: PF04104; DNA primase lrg. 1.
 KW Transferrase; DNA replication; DNA-directed RNA polymerase; Primosome; DNA-binding.
 KW SEQUENCE 509 AA; 58778 MW; B1P500EC442BEP9C CRC64;
 SO Feature Match
 Best Local Similarity 68.2%; Score 30; DB 1; Length 509;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 2 KGTQOY 7
 DB 65 KGTQY 70
 RESULT 17
 ID CH60_ACTAC STANDARD; PRT; 546 AA.
 AC P46398;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 DE GROEL OR GROEL OR MOPA.
 GN Actinobacillus actinomycetemcomitans (Haemophilus actinomycetemcomitans).
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxID=714;
 RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-39.
 RC STRAIN=Y4;
 RX MEDLINE=96017061; PubMed=7567064;
 RA Nakano T., Inai Y., Yamashita Y., Kuzuzaki-Nagira T., Nagaoka S.,
 RA Okahashi N., Koga T., Nishihara T.;
 RT "Molecular and immunological characterization of a 64-kDa protein of Actinobacillus actinomycetemcomitans."
 RL Oral Microbiol. Immunol. 10:151-159 (1995).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and proper assembly of unfolded polypeptides generated under stress conditions (by similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of 7 subunits (by similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D28817; BAA05977.1; -.
 DR HSPD: P06139; IGRL.
 DR HAMAP: MF_06000; -; 1.
 DR InterPro: IPR001844; Chaperin_Cpn60.
 DR InterPro: IPR002423; Cpn60_TCP-1.
 DR Pfam: PF00118; Cpn60_TCP1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 KW Chaperone; ATP-binding.
 KW INIT_MPT 0
 KW SEQUENCE 546 AA; 57312 MW; 6249B0E46334F6F6 CRC64;
 SO Feature Match
 Best Local Similarity 71.4%; Score 30; DB 1; Length 546;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 NKGTQOY 7
 DB 478 NAGTEQY 484
 RESULT 18
 ID CH60_ACTPL STANDARD; PRT; 546 AA.
 AC P94166;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 DE GROEL OR GROEL OR MOPA.
 GN Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxID=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S 4074 / Serotype 1;
 RX MEDLINE=97189570; PubMed=9037757;
 RA Vezina G., Stros M., Clatroux N., Boissinot M.;
 RT "Cloning and characterization of the groE locus from Actinobacillus pleuropneumoniae."
 RL FEMS Microbiol. Lett. 147:11-16 (1997).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and proper assembly of unfolded polypeptides generated under stress conditions (by similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of 7 subunits (by similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).

CC -1- SIMILARITY: Belongs to the chaperronin (HSP60) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U55016; AAB51437.1; -.
CC HSSP: P06139; 1GRU.
CC HAMAP: MF_00600; -; 1.
CC InterPro: IPR001844; Chaparrin Cpn60.
CC InterPro: IPR002423; Cpn60/TCF-1.
CC Pfam: PF00118; cpn60_TCF1; 1.
CC PRINTS: PR00298; CHAPERONIN60.
CC PROSITE: PS00296; TCOMPLEXTCP1.
CC PROSITE: PS00296; CHAPERONINS_CPN60; 1.
CC Chaperrone; ATP-binding.
CC INIT MET 0
SQ SEQUENCE 546 AA; 57513 MW; E84BB72C9BD3DB56 CRC64;

QY Query Match 68.2%; Score 30; DB 1; Length 546;
Best Local Similarity 71.4%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 479 NAGTEQY 485

QY 1 NKGTOQY 7
ID CH60_PASMTU STANDARD; PRT; 547 AA.
AC Q59687;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60 kDa chaperronin (Protein Cpn60) (groEL protein).
GN GROEL OR GROEL OR MOPA OR PM1107.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_Taxid=747;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=P1059;
MEDLINE=96105224; PubMed=8529887;
RA Love B.C., Hansen L.M., Hirsch D.C.;
RT "Cloning and sequence of the groEL heat-shock operon of Pasteurella
RL multocida.";
RN Gene 166:179-180 (1995).
RN [2]
RP SEQUENCE FROM N.A.
STRAIN=Pm70;
RC MEDLINE=21145866; PubMed=11248100;
RA May B.U., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genome sequence of Pasteurella multocida Pm70.";
RC Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
-1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SIMILARITY: Belongs to the chaperronin (HSP60) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U30165; AAA84916.1; -.
CC EMBL: AE006151; AKK03191.1; -.
CC DR PIR: J04519; J04519.
CC HSSP: P06139; 1JON.
CC HAMAP: MF_00600; -; 1.
CC InterPro: IPR001844; Chaparrin Cpn60.
CC InterPro: IPR002423; Cpn60/TCF-1.
CC Pfam: PF00118; cpn60_TCF1; 1.
CC PRINTS: PR00298; CHAPERONIN60.
CC PROSITE: PS00296; TCOMPLEXTCP1.
CC PROSITE: PS00296; CHAPERONINS_CPN60; 1.
CC Chaperrone; ATP-binding; Complete proteome.
CC CONFLICT 424 424 S -> N (IN REF. 1).
CC CONFLICT 464 464 I -> V (IN REF. 1).
SQ SEQUENCE 547 AA; 57291 MW; EBB95B9F2358855E CRC64;

QY Query Match 68.2%; Score 30; DB 1; Length 547;
Best Local Similarity 71.4%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 479 NAGTEQY 485

QY 1 NKGTOQY 7
ID CH60_HAEIN STANDARD; PRT; 548 AA.
AC P43733;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60 kDa chaperronin (Protein Cpn60) (groEL protein).
GN GROEL OR MOPA OR GROEL OR H10543.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=Rd / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RA Fleischmann A.R., Bult C.J., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA McKenney K., Smithey R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Ullrich T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Georgagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RL Rd.";
RC Science 269:496-512 (1995).
-1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the chaperronin (HSP60) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U32736; AAC22201.1; -
 DR PIR; C64076; C64076.
 DR HSSP; P06139; 1JON.
 DR TIGR; H10543; -
 DR HAMAP; MF_00600; -; 1.
 DR InterPro; IPR001844; Chapermin_Cpn60.
 DR InterPro; IPR002423; Cpn60_TCP-1.
 DR Pfam; PF00118; Cpn60_TCP1; 1.
 DR PRINTS; PR00298; CHAPERONIN60.
 DR PRINTS; PR00304; TCOMPLEXTCP1.
 DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
 DR Chaperone; ATP-binding; Complete proteome.
 KM SEQUENCE 548 AA; 57577 MW; CA4066AAC1B62159 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 548;
 Best Local Similarity 71.4%; Pred. No. 74;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGKQOY 7
 |||:
 480 NAGTQOY 486

RESULT 21

AMPB_HUMAN STANDARD; PRT; 650 AA.
 ID AMPB_HUMAN STANDARD; PRT; 650 AA.
 AC Q9H4A4; Q9H4V9; Q9H4D4; Q9NPT7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Aminopeptidase B (EC 3.4.11.6) (Ap-B) (Arginyl aminopeptidase)
 DE (Arginine aminopeptidase).
 GN RNPEP OR APB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Plesse C., Tyms M., Garrafa E., Gouzy C., Lacasa M., Cadel S.,
 RA Foulon T., Cohen P.;
 RT "Human aminopeptidase B on chromosome 1q32.2: cDNA, genomic structure
 RT and expression.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=2238257; PubMed=12477932;
 RA Struhsberg R.L., Colling E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gnarant P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maira M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 291-650 FROM N.A.
 RC TISSUE=Brain;
 RA Bloeker H., Boecker M., Brandt P., Nemes H.-W., Weil B., Wiemann S.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Exopeptidase which selectively removes arginine and/or
 CC lysine residues from the N-terminus of several peptide substrates
 CC including Arg(0)-leu-enkephalin, Arg(0)-Met-enkephalin and Arg(-
 CC 1)-lys(0)-somatostatin-14. Can hydrolyze leucorine A4 (LTA-4)
 CC into leucorine B4 (LTB-4) (By similarity).
 CC -1- CATALYTIC ACTIVITY: Release of a N-terminal Arg and Lys from
 CC oligopeptides when P1' is not Pro. Also acts on 2-naphthylamides
 CC of Arg and Lys.
 CC -1- COFACTOR: BINDS A ZINC ION (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family M1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AJ242586; CAC12957.1; -
 CC EMBL; AJ296161; CAC14047.1; -
 CC EMBL; BC001064; AAH01064.1; -
 CC EMBL; BC012166; AAH12166.1; -
 CC EMBL; AL390139; CAB99087.1; -
 CC PIR; T51870; T51870.
 CC DR PIR; T51870; T51870.
 CC DR Genew; HGNC:10078; RNPEP.
 CC MIM: 602675; -
 DR MEROPS; M01.014; -
 DR InterPro; IPR001930; Ala peptase.
 DR InterPro; IPR006025; Zn Mpeptase.
 DR Pfam; PF01433; Peptidase_M1; 1.
 DR PRINTS; PR00756; ALADIPRASE.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Aminopeptidase; Hydrolase; Zinc; Metalloprotease.
 FT METAL 325 325
 FT ACT SITE 326 326
 FT METAL 329 329
 FT METAL 348 348
 FT ACT SITE 414 414
 FT CONFLICT 2 2
 FT CONFLICT 11 14
 FT CONFLICT 149 149
 FT CONFLICT 153 153
 FT CONFLICT 208 210
 FT CONFLICT 262 262
 FT CONFLICT 262 262
 SQ SEQUENCE 650 AA; 72595 MW; 4C04FB09689F2487 CRC64;
 QY 1 NKGKQOY 8
 |||:
 597 NKGKQOY 604
 Db 597 NKGKQOY 604
 Query Match 68.2%; Score 30; DB 1; Length 650;
 Best Local Similarity 62.5%; Pred. No. 88;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OS Yersinia pestis.
 OG Plasmid pCD1.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 ON NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIM;
 RA Hu P., Elliott J., McCreedy P., Skowronski E., Ganes J.,
 RA Kobayashi A., Carrano A.V., Brubaker R., Garcia E.,
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIMS / Blovax Mediaevalis;
 RX MEDLINE=98427122; PubMed=9746557;
 RA Perry R.D., Straley S.C., Fetherston J.D., Rose D.J., Gregor J.,
 RA Blattner F.R.;
 RT "DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of
 RT Yersinia pestis KIMS";
 RT Infect. Immun. 66:4611-4623(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Blovax Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Tildall R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Fellwell T., Hamlin N., Holtroyd S., Jagsels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail K., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague";
 RT Nature 413:523-527(2001).
 CC -1- FUNCTION: Acts as a virulence determinant.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF053946; AAC62602.1; -;
 CC EMBL; AF074612; AAC69765.1; -;
 CC EMBL; AL117189; CAB54949.1; -;
 CC InterPro: IPR000719; Prot. Kinase.
 CC InterPro: IPR002290; Ser. Thr. kinase.
 CC InterPro: IPR003547; Yers. ser/thr_kin.
 CC Pfam: PF00069; pkinase; 1.
 CC PRINTS: PR01373; YERSSTKINASE.
 CC DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 CC DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 CC DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 CC DR Virulence; Plasmid; Transferase; Serine/threonine-protein kinase;
 CC KW ATP-binding; Signal; Complete proteome.
 CC FT CHAIN 1 732 POTENTIAL.
 CC FT SIGNAL 1 732 PROTEIN KINASE YPKA.
 CC FT DOMAIN 136 408 ATP (BY SIMILARITY).
 CC FT NP BIND 142 150 ATP (BY SIMILARITY).
 CC FT BINDING 163 163 ATP (BY SIMILARITY).
 CC FT ACT SITE 270 270 BY SIMILARITY.
 CC FT CONFLICT 647 647 R -> Q (IN REF. 1 AND 2).
 CC FT ACT SITE 270 270 BY SIMILARITY.
 CC SQ SEQUENCE 732 AA; 81760 MW; DEF803AAE2F5BCD CRC64;
 CC
 CC Query Match 68.2%; Score 30; DB 1; Length 732;
 CC Best Local Similarity 83.3%; Pred. NO. 1e+02; Indels 0; Gaps 0;
 CC Matches 5; Conservative 1; Mismatches 0;

DB 653 GTEQYT 658
 ||:||||
 RESULT 23
 ID YPKA YERPS STANDARD; PRT; 732 AA.
 AC 005608;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein kinase ypkA precursor (EC 2.7.1.37) (Protein kinase A)
 DE (targeted effector protein kinase).
 GN YPKA.
 OS Yersinia pseudotuberculosis.
 OG Plasmid pIB1.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 ON NCBI_TaxID=633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YP111;
 RX MEDLINE=93180911; PubMed=8441468;
 RA Galinov E.E., Haakanson S., Forsberg A., Wolf-Watz H.;
 RT "A secreted protein kinase of Yersinia pseudotuberculosis is an
 RT indispensable virulence determinant";
 RL Nature 361:730-732(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YP111;
 RX MEDLINE=94321324; PubMed=8045884;
 RA Galinov E.E., Haakanson S., Wolf-Watz H.;
 RT "Characterization of the operon encoding the YpkA Ser/Thr protein
 RT kinase and the YopJ protein of Yersinia pseudotuberculosis";
 RT J. Bacteriol. 176:4543-4548(1994).
 CC -1- FUNCTION: Acts as a virulence determinant.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X69439; CAA49215.1; -;
 CC DR EMBL; L33833; AAA68487.1; -;
 CC PIR: S30060; 530060.
 CC DR PhosSite; Q05608; -;
 CC InterPro: IPR000719; Prot. Kinase.
 CC InterPro: IPR002290; Ser. Thr. kinase.
 CC InterPro: IPR003547; Yers. ser/thr_kin.
 CC Pfam: PF00069; pkinase; 1.
 CC PRINTS: PR01373; YERSSTKINASE.
 CC DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 CC DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 CC DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 CC DR Virulence; Plasmid; Transferase; Serine/threonine-protein kinase;
 CC KW ATP-binding; Signal.
 CC FT CHAIN 1 732 POTENTIAL.
 CC FT SIGNAL 1 732 PROTEIN KINASE YPKA.
 CC FT DOMAIN 136 408 ATP (BY SIMILARITY).
 CC FT NP BIND 142 150 ATP (BY SIMILARITY).
 CC FT BINDING 163 163 ATP (BY SIMILARITY).
 CC FT ACT SITE 270 270 BY SIMILARITY.
 CC SQ SEQUENCE 732 AA; 81707 MW; 7C09BDD4F9C9BDF CRC64;
 CC
 CC Query Match 68.2%; Score 30; DB 1; Length 732;
 CC Best Local Similarity 83.3%; Pred. NO. 1e+02; Indels 0; Gaps 0;
 CC Matches 5; Conservative 1; Mismatches 0;

QY 3 GTCQYT 8.
Db 653 GTCQYT 658

RESULT 24

Y506_YEAST STANDARD; PRT; 1121 AA.
ID P32644;
AC P32644;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Hypothetical 127.0 kDa protein in RAD24-BMH1 intergenic region.
GN YER176W OR STGP-ORF61.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
[1]

SEQUENCE FROM N.A.

STRAIN=S288C / AB972;
RX PubMed-916986;
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA Arayo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Huntke-Smith S., Hyman R.W., Kayser A., Komp C., Laskari D., Lew H.,
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petrel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL Nature 387:78-81(1997).
CC -1 SIMILARITY: BELONGS TO THE DNA2/NAM7 HELICASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U18922; AAB64703.1; -
DR PIR; S30862; S30862.
DR SGP; S0000978; YER176W.
DR GO; GO:0005844; C:polyome; IDA.
DR GO; GO:0003678; F:DNA helicase activity; IDA.
DR GO; GO:0006449; F:regulation of translational termination; IMP.
DR InterPro; IPR001410; DEAD.
DR SMART; SMO0487; DEXDc; 1.
KM Hypothetical protein; ATP-binding; Helicase.
NP BIND 670 677
FT NP BIND 1121 AA; 126970 MW; 641CANA6810282A0 CRC64;
SQ SEQUENCE

Query Match 68.2%; Score 30; DB 1; Length 1121;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
Db 382 NKGTSRWT 389

RESULT 25

P1MT_HELPY STANDARD; PRT; 209 AA.
ID P56133;
AC P56133;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-SEP-2003 (Rel. 41, Last sequence update)
DE Protein-L-isoaspartate O-methyltransferase (EC 2.1.1.77) (Protein-
DE beta-aspartate methyltransferase) (P1MT) (Protein L-isoaspartyl
DE methyltransferase) (L-isoaspartyl protein carboxyl methyltransferase).

GN PCM OR HP0363.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
NCBI_TaxID=210;
[1]

SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RC MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khairat H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthy L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).

CC -1 FUNCTION: CATALYZES THE METHYL ESTERIFICATION OF L-ISOASPARTYL
CC RESIDUES IN PEPTIDES AND PROTEINS THAT RESULT FROM SPONTANEOUS
CC DECOMPOSITION OF NORMAL L-ASPARTYL AND L-ASPARAGINYL RESIDUES. IT
CC PLAYS A ROLE IN THE REPAIR AND/OR DEGRADATION OF DAMAGED PROTEINS
CC (BY SIMILARITY).
CC -1 CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-beta-
CC aspartate = S-adenosyl-L-homocysteine + protein L-beta-aspartate
CC methyl ester.
CC -1 SUBUNIT: Monomer (By similarity).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1 SIMILARITY: BELONGS TO THE L-ISOASPARTYL/D-ASPARTYL PROTEIN
CC METHYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AB000553; AAD07431.1; -
DR PIR; C64565; C64565.
DR TIGR; HP0363; -; 1.
DR HAMAP; MF_00090; -; 1.
DR InterPro; IPR000682; Pcmr.
DR InterPro; IPR000051; SAM_bind.
DR Pfam; PF01135; PCMT; 1.
DR TIGRFAMs; TIGR00080; pmt; 1.
DR PROSITE; PS01279; PCMT; 1.
KM Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 209 AA; 23867 MW; 48F060C1C00B50D8 CRC64;

Query Match 65.9%; Score 29; DB 1; Length 209;
Best Local Similarity 71.4%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYT 7
Db 133 NKGMEOYT 139

RESULT 26

UBS1_YEAST STANDARD; PRT; 277 AA.
ID UBS1_YEAST
AC P38290;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Ubiquitin-conjugating enzyme suppressor 1.
GN UBS1 OR YAR165W OR YBR1217.
OS Saccharomyces cerevisiae (Baker's Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxID=4932;
 RN (1)
 RN SEQUENCE FROM N.A.
 RX MEDLINE=9610495; PubMed=8552096;
 RA Prendergast J.A., Ptak C., Kornitzer D., Steussy C.N., Hodgins R.,
 RA Goebel M., Ellison M.J.;
 RT "Identification of a positive regulator of the cell cycle ubiquitin-
 RT conjugating enzyme Cdc34 (Ubc3).";
 RL Mol. Cell. Biol. 16:677-684(1996).
 RN (2)
 RN SEQUENCE FROM N.A.
 RA Entian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,
 RA Niegemann E., Schenk-Groening R., Servos U., Wehner E.,
 RA Wolter R., Brendel M., Bauer J., Braun H., Dorn H., Duesterhus S.,
 RA Gruendlein R., Hedges D., Klebau P., Korol S., Krems B., Proft M.,
 RA Siegers K., Baur A., Boles E., Miesha T.,
 RA Schaeff-Gerstenachlaeger I., Zimmermann F.K.;
 Submitted (AUG-1994) to the EMBL/GenBank/DDJ databases.
 CC -1- FUNCTION: NOT KNOWN, ITS ELEVATED EXPRESSION SUPPRESSES THE
 CC CONDITIONAL CELL CYCLE DEFECTS ASSOCIATED WITH UBC3/CDC34
 CC MUTATIONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S80871; AAB50692.1; -;
 DR EMBL: Z36034; CAAB5126.1; -;
 DR PIR: S46036; S46036.
 DR SGD: S0000369; UBS1.
 DR GO: GO:0005634; C:nucleus; IDA.
 DR GO: GO:0016567; P:protein-nucleus export; IGI.
 DR GO: GO:000611; P:protein-nucleus export; IGI.
 DR SEQ: 277 AA; 32197 MW; 7AB57B0CBABF34B CRC64;
 SQ
 Query Match 65.9%; Score 29; DB 1; Length 277;
 Best Local Similarity 50.0%; Pred. No. 59;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTQOYT 8
 265 NRGSEHYT 272
 RESULT 27
 F16P_RHOSH STANDARD; PRT; 333 AA.
 ID F16P_RHOSH
 AC P27994;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Fructose-1,6-bisphosphatase I (EC 3.1.3.11) (D-fructose-1,6-
 DE bisphosphate 1-phosphohydrolase) (FBPase I).
 GN FBPA.
 OS Rhodospirillum rubrum (Rhodospirillum rubrum).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 OC Rhodospirillaceae; Rhodospirillum.
 OC NCBI_TaxID=1063;
 RN (1)
 RN SEQUENCE FROM N.A.
 RX MEDLINE=91317831; PubMed=1907281;
 RA Gibson J.L., Falcone D.L., Tabita F.R.;
 RT "Nucleotide sequence, transcriptional analysis, and expression of
 RT gene encoded within the form I CO2 fixation operon of Rhodospirillum
 RT rubrum.";
 RL J. Biol. Chem. 266:14646-14653(1991).

CC -1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate + H(2)O = D-
 CC fructose 6-phosphate + phosphate.
 CC -1- ENZYME REGULATION: FRUCTOSE-1,6-BISPHOSPHATASE II IS NOT LIGHT-
 CC ACTIVATED.
 CC -1- PATHWAY: KEY ENZYME OF THE REDUCTIVE PENTOSE PHOSPHATE PATHWAY
 CC OR CALVIN CYCLE OF PHOTOSYNTHETIC CARBON DIOXIDE ASSIMILATION.
 CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR FBPA IN R. RUBRUM.
 CC -1- SIMILARITY: BELONGS TO THE FBPA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M64624; AAA26112.1; -;
 DR PIR: A40767; PAREAS.
 DR HSSP: P46275; IDBZ.
 DR InterPro: IPR000146; In FB_phphatase.
 DR Pfam: PF00316; FBPAse; 1.
 DR PRINTS: PR00377; INFPHPHATASE.
 DR ProDom: PD001491; In FB_phphatase; 1.
 DR PROSITE: PS00124; FBPAse; 1.
 KW Hydrolase; Carbohydrate metabolism; Calvin cycle; Photosynthesis;
 KW ACT_SITE
 FT ACT_SITE 257 BY SIMILARITY.
 SQ SEQUENCE 333 AA; 35996 MW; 7274A0D4C8F8426 CRC64;
 QY 2 KGTQOYT 7
 158 KGTQOYT 163
 Db 158 KGTQOYT 163
 RESULT 28
 SPFB_STRPY STANDARD; PRT; 398 AA.
 ID SPFB_STRPY
 AC P00788; P28296; Q54960; Q54961; Q54962; Q54963; Q54964; Q54965;
 AC Q54966; Q54967; Q54968; Q57024; Q57082; Q57202; Q57211; Q57212;
 AC Q95680;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Streptococcus pyogenes (EC 3.4.22.10) (Streptococcal cysteine
 DE proteinase) (Streptococcus peptidase A) (SPF) (Exotoxin type B) (SPF
 DE B).
 GN SPFB OR SPY2039 OR SPYM3_1742 OR SPY1739 OR SPYM18_2059.
 OS Streptococcus pyogenes, (serotype M3), and
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1314, 198466, 186103;
 RN (1)
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 28-32 AND 146-162.
 RP STRAIN=86-858, and NY-5;
 RX MEDLINE=90330563; PubMed=2198264;
 RA Hauser A.R., Schlievert P.M.;
 RT "Nucleotide sequence of the streptococcal pyrogenic exotoxin type B
 RT gene and relationship between the toxin and the streptococcal
 RT proteinase precursor.";
 RL J. Bacteriol. 172:4536-4542(1990).
 RN (2)
 RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RP STRAIN=789 / Serotype M1, 327 / Serotype M2, 75 / Serotype M3,
 RC 157 / Serotype M3, 315 / Serotype M3, 1251 / Serotype M3,
 RC 321 / Serotype M4, 1289 / Serotype M5, 303 / Serotype M6,
 RC 429 / Serotype M8, 796 / Serotype M9, 800 / Serotype M9,

RC 1896 / Serotype M10, 650 / Serotype M11, 282 / Serotype M12,
 RC 659 / Serotype M13, 660 / Serotype M14, 1898 / Serotype M15,
 RC 1233 / Serotype M17, 156 / Serotype M18, 300 / Serotype M18,
 RC 1294 / Serotype M19, 162 / Serotype M22, 1901 / Serotype M23,
 RC 684 / Serotype M24, 686 / Serotype M25, 1838 / Serotype M27,
 RC 587 / Serotype M28, 366 / Serotype M30, 427 / Serotype M31,
 RC 807 / Serotype M33, 1841 / Serotype M41, 1842 / Serotype M43,
 RC 1226 / Serotype M44, 719 / Serotype M49, 1864 / Serotype M56,
 RC 1882 / Serotype M59, 168 / Serotype M66, 302 / Serotype M73,
 RC 758 / Serotype M75, 1911 / Serotype M75, 1832 / Serotype M76,
 RC 165 / Serotype M, 262 / Serotype M, 317 / Serotype M,
 RC 1990 / Serotype M, 1991 / Serotype M, 2017 / Serotype M,
 RC 2018 / Serotype M, 1719 / Serotype T8, 289 / Serotype T28,
 RC 1590, 1870, 1871, 1872, 1893, and 1914A;
 RC MEDLINE=94285789; PubMed=7516997;
 RA Kapur V., Topouzis S., Majesky M.W., Li L.L., Hamrick M.R.,
 RA Hamill R.J., Patti J.M., Musser J.M.,
 RT "A conserved Streptococcus pyogenes extracellular cysteine protease
 RT cleaves human fibronectin and degrades vitronectin.";
 RT Microb. Pathog. 15:337-346(1993).
 [3]
 RC SEQUENCE FROM N.A.
 RC STRAIN=A-20 / Serotype M1,T1;
 RA Wu J.-J.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 [4]
 RN SEQUENCE FROM N.A.
 RN STRAIN=SV / Serotype M23;
 RC Hong K.;
 RT "A novel cloning method used arbitrarily primed PCR";
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 [5]
 RN SEQUENCE FROM N.A.
 RN STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RC MEDLINE=21192584; PubMed=11296396;
 RA Ferrer J.J., Keshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Perreault C., Seaton S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najaf F.Z., Ren Q., Zhu L., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RT "Complete genome sequence of an M1 strain of Streptococcus
 RT pyogenes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 [6]
 RN SEQUENCE FROM N.A.
 RN STRAIN=MGAS315 / Serotype M3;
 RC MEDLINE=22133808; PubMed=12122206;
 RA Betes S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.;
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 RT emergence.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 [7]
 RN SEQUENCE FROM N.A.
 RN STRAIN=SSI-1 / Serotype M3;
 RC Nakagawa I., Kurakawa K., Nakata M., Tomiyasu Y., Yamashita A.,
 RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
 RA Hayaishi H., Hamada S.;
 RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
 RT of S. pyogenes SSI-1, SF370 and MGAS832.";
 RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 [8]
 RN SEQUENCE FROM N.A.
 RN STRAIN=MGAS8232 / Serotype M18;
 RC MEDLINE=21927593; PubMed=11917108;
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Sturdevant D.B., Ricklefs S.M., Porcella S.F.,
 RA Parkins L.D., Betes S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18
 RT group A Streptococcus strains associated with acute rheumatic fever

RT outbreaks.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 [9]
 RN PRELIMINARY SEQUENCE OF 28-86 AND 121-398.
 RA Yonaha K., Elliott S.D., Liu T.-Y.;
 RT "Primary structure of zymogen of streptococcal proteinase.";
 RT J. Protein Chem. 1:317-334(1982).
 [10]
 RN PRELIMINARY SEQUENCE OF 146-398.
 RA Tai J.-Y., Kott A.A., Liu T.-Y., Elliott S.D.;
 RT "Primary structure of streptococcal proteinase. III. Isolation of
 RT cyanogen bromide peptides: complete covalent structure of the
 RT polypeptide chain.";
 RT J. Biol. Chem. 251:1955-1959(1976).
 [11]
 RN FUNCTION.
 RC STRAIN=N2131 / Serotype M49,T14;
 RA MEDLINE=99081733; PubMed=9864206;
 RA Kuo C.-F., Wu J.-J., Tsai P.-J., Kao F.-J., Lei H.-Y., Lin M.T.,
 RA Lin Y.-S.;
 RT "Streptococcal pyrogenic exotoxin B induces apoptosis and reduces
 RT phagocytic activity in U937 cells.";
 RT Infect. Immun. 67:126-130(1999).
 [12]
 RN FUNCTION.
 RC STRAIN=A-20 / Serotype M1,T1;
 RA MEDLINE=99386917; PubMed=10456871;
 RA Tsai P.-J., Lin Y.-S., Kuo C.-F., Lei H.-Y., Wu J.-J.;
 RT "Group A Streptococcus induces apoptosis in human epithelial cells.";
 RT Infect. Immun. 67:4334-4339(1999).
 CC -1- FUNCTION: IMPORTANT STREPTOCOCCAL VIRULENCE FACTOR WHICH CLEAVES
 CC HUMAN FIBRONECTIN AND DEGRADES VITRONECTIN. ALSO CLEAVES HUMAN
 CC IL1B PRECURSOR TO FORM BIOLOGICALLY ACTIVE IL1B. CAN INDUCE
 CC APOPTOSIS IN HUMAN MONOCYTES AND EPITHELIAL CELLS IN VITRO, AND
 CC REDUCES PHAGOCYTIC ACTIVITY IN MONOCYTIC CELLS. THUS, MAY PLAY A
 CC ROLE IN BACTERIAL COLONIZATION, INVASION, AND INHIBITION OF WOUND
 CC HEALING.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage with hydrophobic
 CC residues at P2, P1 and P1'.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: THE STREPTOCOCCAL PYROGENIC TOXINS A, B, AND C ARE
 CC THE CAUSATIVE AGENTS OF THE SYMPTOMS ASSOCIATED WITH SCARLET
 CC FEVER, HAVE BEEN ASSOCIATED WITH STREPTOCOCCAL TOXIC SHOCK-LIKE
 CC DISEASE AND MAY PLAY A ROLE IN THE EARLY EVENTS OF RHEUMATIC
 CC FEVER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C10.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, M86905; AAA26978.1; -
 DR EMBL, L26125; AAA26879.1; -
 DR EMBL, L26126; AAA26992.1; -
 DR EMBL, L26127; AAA26993.1; -
 DR EMBL, L26128; AAA26994.1; -
 DR EMBL, L26130; AAA26996.1; -
 DR EMBL, L26131; AAA26997.1; -
 DR EMBL, L26132; AAA26998.1; -
 DR EMBL, L26133; AAA26999.1; -
 DR EMBL, L26134; AAA27000.1; -
 DR EMBL, L26135; AAA27001.1; -
 DR EMBL, L26136; AAA27002.1; -
 DR EMBL, L26137; AAA27003.1; -
 DR EMBL, L26138; AAA27004.1; -
 DR EMBL, L26139; AAA27005.1; -
 DR EMBL, L26140; AAA27006.1; -
 DR EMBL, L26141; AAA27007.1; -

DR EMBL, L26142; AAA27008.1; -
DR EMBL, L26143; AAA27009.1; -
DR EMBL, L26144; AAA27010.1; -
DR EMBL, L26145; AAA27011.1; -
DR EMBL, L26146; AAA27012.1; -
DR EMBL, L26147; AAA27013.1; -
DR EMBL, L26148; AAA27014.1; -
DR EMBL, L26149; AAA27015.1; -
DR EMBL, L26150; AAA27016.1; -
DR EMBL, L26151; AAA26980.1; -
DR EMBL, L26152; AAA26981.1; -
DR EMBL, L26153; AAA26982.1; -
DR EMBL, L26154; AAA26983.1; -
DR EMBL, L26155; AAA26984.1; -
DR EMBL, L26156; AAA26985.1; -
DR EMBL, L26157; AAA26986.1; -
DR EMBL, L26158; AAA26987.1; -
DR EMBL, L26159; AAA26988.1; -
DR EMBL, L26160; AAA26989.1; -
DR EMBL, L26161; AAA26990.1; -
DR EMBL, L26162; AAA26991.1; -
EMBL, AF104940; AAD17930.1; -.

CC	-1-	CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC		alcohol + phosphate.
CC	-1-	COFACTOR: IRON AND ZINC.
CC	-1-	SUBUNIT: Homodimer; disulfide-linked.
CC	-1-	SUBCELLULAR LOCATION: Secreted.
CC	-1-	PTM: THE C-TERMINUS OF THE ENZYME IS OFTEN MODIFIED BY THE
CC		REMOVAL OF FOUR TERMINAL RESIDUES.
DR	PDB; 1KBP; 14-OCT-96.	
DR	PDB; 3KBP; 07-DEC-96.	
DR	PDB; 4KBP; 07-DEC-96.	
DR	GlycoSuiteDB; P80366; -	
DR	InterPro; IPR004843; M-peptidase.	
DR	Pfam; PF00149; Metallopross; 1.	
KW	Hydrolase; Iron; Zinc; Metal-binding; Glycoprotein; 3D-structure.	
FT	METAL	135 135 IRON.
FT	METAL	164 164 ZINC AND IRON.
FT	METAL	167 167 IRON.
FT	METAL	201 201 ZINC.
FT	ACT SITE	202 202 GENERAL ACID (BY SIMILARITY).
FT	METAL	286 286 ZINC.
FT	METAL	323 323 ZINC.
FT	METAL	325 325 IRON.
FT	DISULFID	345 345 INTERCHAIN.
FT	CARBOHYD	81 81 N-LINKED (GLCNAC. . .) (PARTIAL).
FT	CARBOHYD	109 109 N-LINKED (GLCNAC. . .)
FT	CARBOHYD	143 143 N-LINKED (GLCNAC. . .)
FT	CARBOHYD	211 211 N-LINKED (GLCNAC. . .)
FT	CARBOHYD	396 396 N-LINKED (GLCNAC. . .)

RT "The amino acid sequence of the red kidney bean Fe(III)-Zn(II) purple
RT acid phosphatase. Determination of the amino acid sequence by a
RT combination of matrix-assisted laser desorption/ionization mass
RT spectrometry and automated Edman sequencing.";
RL Eur. J. Biochem. 228:369-375(1994).
RN [2]
RP PARTIAL SEQUENCE, AND STRUCTURE OF CARBOHYDRATES.
RC TISSUE=Seed;
RC MEDLINE=94110779; PubMed=8125089;
RX Stehl B., Klabunde T., Witzel H., Krebs B., Steup M., Karas M.,
RA Hillel-Berg F.;
RA Klabunde T., Witzel H., Krebs B., Steup M., Karas M.,
RT "The oligosaccharides of the Fe(III)-Zn(II) purple acid phosphatase
RT of the red kidney bean. Determination of the structure by a
RT combination of matrix-assisted laser desorption/ionization mass
RT spectrometry and selective enzymic degradation.";
RL Eur. J. Biochem. 220:321-330(1994).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS).
RX MEDLINE=96275658; PubMed=8688579.
RA Klabunde T., Streeter N., Froehlich R., Witzel H., Krebs B.;
RT "Mechanism of Fe(III)-Zn(II) purple acid phosphatase based on crystal
RT structures.";
RL J. Mol. Biol. 259:737-748(1996)

FT TURN 295 300
 FT HELIX 301 313
 FT SEQUENCE 314 315
 FT STRAND 318 321
 FT STRAND 327 330
 FT STRAND 333 334
 FT STRAND 348 349
 FT TURN 351 352
 FT STRAND 355 359
 FT TURN 364 366
 FT STRAND 372 372
 FT TURN 378 379
 FT STRAND 380 384
 FT STRAND 388 394
 FT STRAND 399 406
 FT TURN 411 412
 FT STRAND 414 421
 FT TURN 423 425
 SQ SEQUENCE 432 AA; 50236 MW; D2A60BCD9385A886 CRC64;

Query Match 65.9%; Score 29; DB 1; Length 432;
 Best Local Similarity 71.4%; Pred. No. 95;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQOYT 8
 Db 258 KGTPOYT 264

RESULT 30
 IDH_THETH STANDARD; PRT; 495 AA.

AC P33197;
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Isocitrate dehydrogenase [NADP] (EC 1.1.1.42) (Oxalosuccinate decarboxylase) (IDH) (NADP+-specific IDH) (IDP).
 GN ICD.

OS Thermus thermophilus.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 OC Thermus.

OX NCBI_TaxID=274;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HB8 / ATCC 27634;
 RX MEDLINE=92171520; PubMed=1539996;

RA Miyazaki K., Egnuchi H., Yamagishi A., Wakagi T., Oshima T.;
 "Molecular cloning of the isocitrate dehydrogenase gene of an extreme thermophile, Thermus thermophilus HB8."

RA Appl. Environ. Microbiol. 58:93-98(1992).

CC -1- CATALYTIC ACTIVITY: Isocitrate + NADP(+) = 2-oxoglutarate + CO(2)
 + NADPH.

CC -1- SUBUNIT: Homodimer.
 -1- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE
 DEHYDROGENASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M94317; AAA27492.1; -

DR HSSP; P39126; IHOS.

DR InterPro; IPR001804; Isocdh.

DR Pfam; PF00180; Isocdh. 1.

DR PROSITE; PS00470; IDH_IMDH; 1.

KW Oxidoreductase; NADP; Glyoxylate bypass; Tricarboxylic acid cycle;
 KW Phosphorylation.

FT ACT_SITE 98 98 BINDING TO ISOCITRATE (BY SIMILARITY).

FT MOD RES 98 98 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 495 AA; 54321 MW; 58D30F239A8DE7BD CRC64;

Query Match 65.9%; Score 29; DB 1; Length 495;
 Best Local Similarity 71.4%; Pred. No. 11e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOY 7
 Db 424 NKGTQOY 430

RESULT 31
 IDH_THETH STANDARD; PRT; 504 AA.

AC O58205;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MJ0795.

GN MJ0795.

OC Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.

OX NCBI_TaxID=2190;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kierulff A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Funtmann J.L., Nguyen D.,
 RA Utermack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Huzar M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii";
 Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RL Science 273:1058-1073(1996).

CC -1- SIMILARITY: TO M. JANNASCHII MJ1506 AND MJ1561.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U67524; AAB98798.1; -

DR FIR; C64399; C64399.

DR TIGR; MJ0795; -

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 6 26 POTENTIAL.

SQ SEQUENCE 504 AA; 57146 MW; A1A659DB7BAEC74B CRC64;

Query Match 65.9%; Score 29; DB 1; Length 504;
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQOYT 8
 Db 117 KGTQOYT 123

RESULT 32
 AARB_AARBPR STANDARD; PRT; 527 AA.

AC O06077; P81374;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Abriin-b precursor [Contains: Abriin-b A chain (rRNA N-glycosidase)
 (EC 3.2.22); Abriin-b B chain].
 OS Abrus precatorius (Indian licorice) (Crab's eye)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 OX NCBI_TaxID=3816;
 (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93132798; PubMed=8421313;
 RT Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
 RT "Primary structure of three distinct isoabrinins determined by cDNA
 RT sequencing. Conservation and significance.";
 RL J. Mol. Biol. 229:263-267(1993).
 RN [2]
 RP SEQUENCE OF 260-527.
 RC TISSUE=Seed.
 RX MEDLINE=93169023; PubMed=7763422;
 RA Kimura M., Sumizawa T., Funatsu G.;
 RA "The complete amino acid sequences of the B-chains of abrin-a and
 RA abrin-b, toxic proteins from the seeds of Abrus precatorius.";
 RA Bioest. Biotechnol. Biochem. 57:166-169(1993).
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
 CC ABRIN-A IS MORE TOXIC THAN RICIN.
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
 CC PRECEDES ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, M89345; AAA32625.1; --
 CC PIR, S32430; S32430.
 CC HSSP, P1140; 1ABR.
 CC InterPro, IPR000772; Ricin_B_lectin.
 CC InterPro, IPR001574; RIP.
 CC Pfam, PF00652; Ricin_B_lectin; 6.
 CC Pfam, PF00161; RIP; 1.
 CC PRINTS, PR00396; SHIGARICIN.
 CC SMART, SM00458; RICIN; 2.
 CC PROSITE, PS00231; RICIN B LECTIN; 2.
 CC PROSITE, PS00275; SHIGA RICIN; 1.
 CC plant defense, Hydrolase, Protein synthesis inhibitor; Toxin; Repeat;
 CC Glycoprotein; Lectin; Pyroglutamate carboxylic acid.
 CC KW Glycoprotein; Lectin; Pyroglutamate carboxylic acid.
 CC -----
 CC CHAIN 1 250
 CC PEPTIDE 251 260
 CC CHAIN 261 527
 CC DOMAIN 272 399
 CC DOMAIN 402 526
 CC REPEAT 282 324
 CC REPEAT 325 365
 CC REPEAT 368 400
 CC REPEAT 413 448
 CC REPEAT 452 491
 CC REPEAT 494 527
 CC ACT SITE 163
 CC DISULFID 246
 CC DISULFID 285 304

FT DISULFID 328 345 BY SIMILARITY.
 FT DISULFID 416 429 BY SIMILARITY.
 FT DISULFID 455 472 BY SIMILARITY.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
 FT CARBOHYD 110 110 SIMILARITY).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 291 291 N -> D (IN REF. 2).
 FT CONFLICT 350 351 D -> N (IN REF. 2).
 FT CONFLICT 378 378 AB -> PQ (IN REF. 2).
 FT CONFLICT 426 426 S -> N (IN REF. 2).
 FT CONFLICT 428 428 L -> M (IN REF. 2).
 FT CONFLICT 431 431 Y -> D (IN REF. 2).
 FT CONFLICT 431 431 N -> S (IN REF. 2).
 FT CONFLICT 484 484 R -> K (IN REF. 2).
 FT CONFLICT 491 491 N -> S (IN REF. 2).
 FT CONFLICT 493 493 H -> Y (IN REF. 2).
 FT CONFLICT 502 502 R -> Q (IN REF. 2).
 FT CONFLICT 509 509 E -> Q (IN REF. 2).
 FT CONFLICT 513 513 H -> W (IN REF. 2).
 FT CONFLICT 516 516 H -> T (IN REF. 2).
 SQ SEQUENCE 527 AA; 59114 MW; 3253AB90CE9494A CRC64;
 Query Match 65.9%; Score 29; DB 1; Length 527;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GTQQT 8
 DB 101 GTQQYS 106
 RESULT 33
 ID GT66 SCHPO STANDARD; PRT; 535 AA.
 AC 074849; Q96TK9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE High-affinity fructose transporter ght6 (Hexose transporter 6)
 GN GHT6 OR MEU12 OR SPC1235.13.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=972;
 RX MEDLINE=20200352; PubMed=10735857;
 RA Heiland S., Radovanovic N., Hoefer M., Winderrickx J., Lichtenberg H.;
 RA "Multiple hexose transporters of Schizosaccharomycetes pombe.";
 RL J. Bacteriol. 182:2153-2162(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle B.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grynolprez B.,
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritze C., Holzer E., Meestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purrelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaur V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt K., Moore K., Hurst S.M.,
 RA Lucas M., Kochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Usery D., Barrett B.G., Nurse P.,
 RT "The genome sequence of *Schizosaccharomyces pombe*."
 RL Nature 415:871-880(2002).
 [3]
 RN SEQUENCE OF 478-535 FROM N.A.
 RP STRAIN=CD16-1;
 RX MEDLINE=21270454; PubMed=11376151;
 RA Watanabe T., Miyahita K., Saito T.T., Yoneki T., Kakhara Y.,
 RA Nabeshima K., Kishi Y.A., Shimoda C., Nojima H.,
 RT "Comprehensive isolation of meiosis-specific genes identifies novel
 proteins and unusual non-coding transcripts in *Schizosaccharomyces
 pombe*."
 CC Nucleic Acids Res. 29:2327-2337(2001).
 CC -1- FUNCTION: High-affinity fructose transporter.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF098076; AAC64976.1; -
 CC EMBL: AL031764; CA21117.1; -
 CC EMBL: AB054532; BAB60886.1; -
 CC PIR: T40887; T40887.
 DR GeneDB: SPombe: SPC1235.13; -
 DR InterPro: IPR007114; MFS.
 DR InterPro: IPR005828; SUB_transporter.
 DR InterPro: IPR005829; Sug_transporter.
 DR InterPro: IPR003663; Sugar_transp.
 DR Pfam: PF00083; sugar_tr; 1.
 DR PRINTS: PR00171; SUGRTRANSPORT.
 DR TIGRfam: TIGR00879; SP. 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
 DR Transmembrane: Sugar transport; Transport; Glycoprotein; Meiosis.
 FT DOMAIN 1 30 9
 FT TRANSMEM 10 58
 FT DOMAIN 31 58
 FT TRANSMEM 59 79
 FT DOMAIN 80 87
 FT TRANSMEM 88 108
 FT DOMAIN 109 112
 FT TRANSMEM 113 133
 FT DOMAIN 134 144
 FT TRANSMEM 145 165
 FT DOMAIN 166 181
 FT TRANSMEM 182 202
 FT DOMAIN 203 268
 FT TRANSMEM 269 287
 FT DOMAIN 288 303
 FT TRANSMEM 304 324
 FT DOMAIN 325 330
 FT TRANSMEM 331 351
 FT DOMAIN 352 365
 FT TRANSMEM 366 386
 FT DOMAIN 387 406
 FT TRANSMEM 407 427
 FT DOMAIN 428 435
 FT TRANSMEM 435 455

FT DOMAIN 456 535 CYTOPLASMIC (POTENTIAL).
 FT CAROHWD 359 359 N-LINKED (GLCNAC...) (POTENTIAL).
 SO SEQUENCE 535 AA; 59403 MW; FECD99A4E51844 CRC64;
 73
 Query Match 65.9%; Score 29; DB 1; Length 535;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NKGTOOYT 8
 Db 166 NMGTHKTY 173
 RESULT 34
 ID FUT2_ARATH STANDARD; PRT; 539 AA.
 AC 081053;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable fucosyltransferase 2 (EC 2.4.1.-) (AtFUT2).
 GN FUT2 OR AT2G03210 OR T18B12.12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldhlym T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Roming C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.D., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.,
 RT "Sequence and analysis of chromosome 2 of the plant *Arabidopsis
 thaliana*."
 RL Nature 402:761-768(1999).
 RN [2]
 RP IDENTIFICATION AS PUTATIVE FUCOSYLTRANSFERASE, AND TISSUE SPECIFICITY.
 RX MEDLINE=21608393; PubMed=11743104.
 RA Sarría R., Wagner T.A., O'Neill M.A., Paik A., Wilkerson C.G.,
 RA Keegstra K., Raikhel N.V.,
 RT "Characterization of a family of Arabidopsis genes related to
 xyloglucan fucosyltransferase1."
 RL Plant Physiol. 127:1595-1606(2001).
 CC -1- FUNCTION: May be involved in cell wall biosynthesis. May
 act as a fucosyltransferase.
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
 form in trans cisternae of Golgi (By similarity).
 CC -1- TISSUE SPECIFICITY: Expressed in roots, stems, leaves, flowers,
 siliques and seedlings.
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 37.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AC005313; AAC34481.1; -
 CC PIR: T02705; T02705.
 DR InterPro: IPR004938; XG_FTase.
 DR Pfam: PF03254; XG_FTase; 1.
 DR Transfaser: Glycosyltransferase; Transmembrane; Glycoprotein;
 KW Signal-anchor; Golgi stack; Cell wall.

```

FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 6 26 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT 27 539 (POTENTIAL).
FT DOMAIN 27 539 LIMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 539 AA; 61570 MW; 5AEAVF45847DC67C CRC64;

Query Match 65.9%; Score 29; DB 1; Length 539;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GTQOYT 8
DB 121 GTRQYT 126

RESULT 35
MC 1 SCHPO STANDARD; PRT; 580 AA.
MC 094547;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Serine/threonine-protein kinase srk1 (EC 2.7.1.37) (Sty1-regulated kinase 1).
GN SRK1 OR SPCCL322.08.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxId=4896;
RN 11
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND
RP PHOSPHORYLATION.
RX MEDLINE=22194329; PubMed=12080074;
RA Smith D.A., Toone W.M., Chen D., Baehler J., Jones N., Morgan B.A.,
RA Quinn J.;
RT "The srk1 protein kinase is a target for the sty1 stress-activated
RL MARK in fission yeast.";
RL J. Biol. Chem. 277:33411-33421(2002).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Baaham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,
RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Glynnmeyer B.,
RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer M., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreesen S., Gloux S., Lelaire V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armerstrong J., Forsberg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovskiy G.V., Usary D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RA Nature 415:871-880(2002).
RL 1- FUNCTION: Has a role in the regulation of meiosis via the sty1

```

```

CC stress-activated MAPK pathway. Inhibits nitrogen-limitation induced
CC arrest at G1.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic, and nuclear.
CC -1- PTM: Phosphorylated by sty1.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
CC EMBL; AL035259; CAA22861.1; -.
CC PIR; T40939; T40939.
CC HSSP; Q63450; 1A06.
CC GeneDB; Spombe; SPCCL322.08; -.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 2.
CC SMART; SM00220; S_TKC; 1.
CC DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
CC DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
CC DR Transferase; Serine/threonine-protein kinase; Meiosis; ATP-binding;
CC KW Nuclear protein; Phosphorylation.
CC FT DOMAIN 124 421 PROTEIN KINASE.
CC FT NP BIND 130 138 ATP (BY SIMILARITY).
CC FT BINDING 153 153 ATP (BY SIMILARITY).
CC FT ACT SITE 257 257 BY SIMILARITY.
CC SQ SEQUENCE 580 AA; 66136 MW; 85BA8DA3C0EDP05B CRC64;

Query Match 65.9%; Score 29; DB 1; Length 580;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KETQYT 8
DB 119 KGEQYT 125

RESULT 36
ID PALY HELAN STANDARD; PRT; 667 AA.
AC 004058;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
GN PAL.
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Asterales; Asteraceae; Asteroideae;
OC Heliantheae; Helianthus.
OC NCBI_TaxId=4232;
OX 11
RP SEQUENCE FROM N.A.
RA Mazyer F.A., Sallès S., Drevet J., Roedel-Drevet P.,
RA Tournelle D., Ledoit G.;
RL (in) Plant Gene Register PGR98-108.
CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),

```

WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
OF RESIDUES ALA-SER-GLY (BY SIMILARITY).

-1- SIMILARITY: Belongs to the PAL / histidase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: Y12461; CAA73065.1; -
CC PIR: T12749; T12749.
CC HSSP: P21310; 1B8F.
CC InterPro: IPR001106; Phe/His NH3lyase.
CC InterPro: IPR005922; Phe_am_lyase.
CC Pfam: PF00221; PAL; 1.
CC TIGRFAMs: TIGR01226; phe_am_lyase; 1.
CC PROSITE: PS00488; PAL_HISTIDASE; 1.
CC Lyase; Phenylpropanoid metabolism; Multigene family.
CC SITE 197 199 MODIFIED TO FORM 4-METHYLLIDENE-IMIDAZOLE-
CC FT SEQUENCE 667 AA; 72076 MW; FBD3DB590E70AB2 CRC64;
CC 5-ONE (BY SIMILARITY).

SQ Query Match 65.9%; Score 29; DB 1; Length 667;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
| | | | |
| | | | |
Db 97 NKGTDYSY 103

RESULT 37
PAL2_LITER STANDARD; PRT; 705 AA.
ID PAL2_LITER
AC O49836;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DS Phenylalanine ammonia-lyase 2 (EC 4.3.1.5) (PAL-2).
OS Lithospermum erythrorhizon.
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Boraginaceae; Lithospermum.
OC NCBI_Taxid=34254;
OX
[1]
SEQUENCE FROM N.A.
MEDLINE=98101967; PubMed=9438980;
RA Yazaki K., Kataoka M., Honda G., Severin K., Heide L.;
RT "cDNA cloning and gene expression of phenylalanine ammonia-lyase in
RT Lithospermum erythrorhizon.";
RL Biochem. Biotechnol. Biochem. 61:1995-2003(1997).
CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN ROOTS.
CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLLIDENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC -1- SIMILARITY: Belongs to the PAL / histidase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: D83076; BAA24929.1; -
CC DR JCS873; JCS873.
CC DR HSSP: P21310; 1B8F.
CC InterPro: IPR001106; Phe/His NH3lyase.
CC InterPro: IPR005922; Phe_am_lyase.
CC Pfam: PF00221; PAL; 1.
CC TIGRFAMs: TIGR01226; phe_am_lyase; 1.
CC PROSITE: PS00488; PAL_HISTIDASE; 1.
CC Lyase; Phenylpropanoid metabolism; Multigene family.
CC SITE 191 193 MODIFIED TO FORM 4-METHYLLIDENE-IMIDAZOLE-
CC FT SEQUENCE 705 AA; 76509 MW; ABBEBF62BC7DB9D CRC64;
CC 5-ONE (BY SIMILARITY).

SQ Query Match 65.9%; Score 29; DB 1; Length 705;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOOY 7
| | | | |
| | | | |
Db 91 NKGTDYSY 97

RESULT 38
PAL1_IPOBA STANDARD; PRT; 707 AA.
ID PAL1_IPOBA
AC P14166;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DS Phenylalanine ammonia-lyase (EC 4.3.1.5).
OS PAL.
OS Ipomoea batatas (Sweet potato) (Batace).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Convolvulaceae; Ipomoea.
OC NCBI_Taxid=4120;
OX
[1]
SEQUENCE FROM N.A.
RA Tanaka Y., Matsuka M., Yamanoto N., Ohashi Y., Kano-Murakami Y.,
RA Ozeki Y.;
RT "Structure and characterization of a cDNA clone for phenylalanine
RT ammonia-lyase from cut-injured roots of sweet potato.";
RL Plant Physiol. 90:1403-1407(1989).
CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLLIDENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC -1- SIMILARITY: Belongs to the PAL / histidase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M29232; AAA3389.1; -
CC PIR: S29029; S29029.
CC HSSP: P21310; 1B8F.
CC InterPro: IPR001106; Phe/His NH3lyase.
CC InterPro: IPR005922; Phe_am_lyase.
CC Pfam: PF00221; PAL; 1.
CC TIGRFAMs: TIGR01226; phe_am_lyase; 1.
CC PROSITE: PS00488; PAL_HISTIDASE; 1.
CC Lyase; Phenylpropanoid metabolism; Multigene family.
CC SITE 192 194 MODIFIED TO FORM 4-METHYLLIDENE-IMIDAZOLE-

FT 5-ONE (BY SIMILARITY).
 SQ SEQUENCE 707 AA; 77138 MW; 8789513A9128F8B CRC64;
 Query Match 65.9%; Score 29; DB 1; Length 707;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
 |||||
 DB 92 NKGTDSY 98

RESULT 39
 ID PAL1 DAUCA STANDARD; PRT; 708 AA.
 AC 023855;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 RT Phenylalanine ammonia-lyase 1 (EC 4.3.1.5).
 RT PAL1.
 RT Daucus carota (Carrot).
 RT Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 RT OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 RT OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
 RT NCBI_TaxId=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=ev. Kurodagosun;
 RX MEDLINE=97478956; PubMed=9337617;
 RA Takeda J., Ozeki Y., Yoshida K.;
 RT "Action spectrum for induction of promoter activity of phenylalanine
 ammonia-lyase gene by UV in carrot suspension cells.";
 RT Photochem. Photobiol. 66:464-470(1997).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the PAL / histidase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL; D85850; BAA23367.1; -.
 CC DR PIR; T14295; T14295.
 CC DR HSSP; P21310; 1B8F.
 CC DR InterPro; IPR001106; Phe/His NH3lyase.
 CC DR InterPro; IPR005922; Phe_am_lyase.
 CC DR Pfam; PF00221; PAL; 1.
 CC DR TIGRPFAM; TIGR01226; phe_am_lyase; 1.
 CC DR PROSITE; PS00488; PAL_HISTIDASE; 1.
 CC KW Lyase; Phenylpropanoid metabolism; Multigene family.
 CC FT SITE 194 196 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-
 CC 5-ONE (BY SIMILARITY).
 FT SQ SEQUENCE 708 AA; 76845 MW; 06104113735E80E3 CRC64;
 Query Match 65.9%; Score 29; DB 1; Length 708;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
 |||||
 DB 94 NKGTDSY 100

RESULT 40
 ID PAL1 LITER STANDARD; PRT; 710 AA.
 AC 049635;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 RT Phenylalanine ammonia-lyase 1 (EC 4.3.1.5) (PAL-1).
 RT Litospermum erythrorhizon.
 RT OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 RT OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 RT OC Asteridae; lamiales; Boraginaceae; Litospermum.
 RT NCBI_TaxId=34254;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=98101967; PubMed=9438980;
 RA Yasaki K., Kataoka M., Honda G., Severin K., Heide L.;
 RT "cDNA cloning and gene expression of phenylalanine ammonia-lyase in
 RT Litospermum erythrorhizon.";
 RT Biosci. Biotechnol. Biochem. 61:1995-2003(1997).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN ROOTS.
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the PAL / histidase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL; D83075; BAA24928.1; -.
 CC DR PIR; JC5872; JC5872.
 CC DR HSSP; P21310; 1B8F.
 CC DR InterPro; IPR001106; Phe/His NH3lyase.
 CC DR InterPro; IPR005922; Phe_am_lyase.
 CC DR Pfam; PF00221; PAL; 1.
 CC DR TIGRPFAM; TIGR01226; phe_am_lyase; 1.
 CC DR PROSITE; PS00488; PAL_HISTIDASE; 1.
 CC KW Lyase; Phenylpropanoid metabolism; Multigene family.
 CC FT SITE 196 198 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-
 CC 5-ONE (BY SIMILARITY).
 FT SQ SEQUENCE 710 AA; 77518 MW; CEFD1D4B3A80ED03 CRC64;
 Query Match 65.9%; Score 29; DB 1; Length 710;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOY 7
 |||||
 DB 96 NKGTDSY 102

Search completed: August 20, 2003, 09:30:43
 Job time : 5.61856 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 23.2577 Seconds
(without alignments)
88.763 Million cell updates/sec

Title: US-09-991-433-6
Perfect score: 44
Sequence: 1 NKGTQOYT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	138	12	085155 human parvo
2	44	100.0	141	12	085171 human parvo
3	44	100.0	141	12	085166 human parvo
4	44	100.0	141	12	085146 human parvo
5	44	100.0	142	12	085168 human parvo
6	44	100.0	142	12	085173 human parvo
7	44	100.0	144	12	085138 human parvo
8	44	100.0	145	12	085181 human parvo
9	44	100.0	145	12	085161 human parvo
10	44	100.0	146	12	085158 human parvo
11	44	100.0	147	12	085142 human parvo
12	44	100.0	148	12	085150 human parvo
13	44	100.0	151	12	085131 human parvo
14	44	100.0	151	12	085177 human parvo
15	44	100.0	153	12	085123 human parvo
16	44	100.0	157	12	085196 human parvo

17	44	100.0	162	12	085135 human parvo
18	44	100.0	546	12	0913X0 human parvo
19	44	100.0	546	12	0913W7 human parvo
20	44	100.0	554	12	09P2S9 human parvo
21	44	100.0	554	12	090201 human parvo
22	44	100.0	554	12	065790 human parvo
23	44	100.0	554	12	0912B7 human eryth
24	44	100.0	554	12	09JGP7 human parvo
25	44	100.0	554	12	09WKL9 human parvo
26	44	100.0	554	12	08JNS4 human parvo
27	44	100.0	760	12	09P2T8 human parvo
28	44	100.0	761	12	09P2U0 human parvo
29	44	100.0	765	12	09P2T6 human parvo
30	44	100.0	769	12	09P2T4 human parvo
31	44	100.0	773	12	0913X1 human parvo
32	44	100.0	773	12	0913W8 human parvo
33	44	100.0	781	12	08JNS6 human parvo
34	44	100.0	781	12	065789 human parvo
35	44	100.0	781	12	090223 human parvo
36	44	100.0	781	12	085191 human parvo
37	44	100.0	781	12	090221 human parvo
38	44	100.0	781	12	089318 human parvo
39	44	100.0	781	12	08JYB3 erythroviru
40	44	100.0	781	12	090224 human parvo
41	44	100.0	781	12	09P2T0 human parvo
42	44	100.0	781	12	089316 human parvo
43	44	100.0	781	12	089320 human parvo
44	44	100.0	781	12	085117 human parvo
45	44	100.0	781	12	089321 human parvo

ALIGNMENTS

```

RESULT 1
ID 085155 PRELIMINARY; PRT; 138 AA.
AC 085155;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Pöblitzki A., Giegler A., Cassinotti P., Siegl G.,
   Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70564; CA94477.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15237 MW; 5899FB8879A3B6BD CRC64;
QY 1 NKGTQOYT 8
Db 107 NKGTQOYT 114

RESULT 2
ID 085171 PRELIMINARY; PRT; 141 AA.
AC 085171;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Pöblotzki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
  "XXXSequence variability among different parvovirus B19 isolates.";
  Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; Z70580; CAA94493.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 141
FT NON_TER 1 141
SQ SEQUENCE 141 AA; 15770 MW; C9E92572A78C6C27 CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 141;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
DB 109 NKGTQOYT 116

RESULT 3
ID 085166 PRELIMINARY; PRT; 141 AA.
AC 085166;
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Pöblotzki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
  "XXXSequence variability among different parvovirus B19 isolates.";
  Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; Z70575; CAA94488.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 141
FT NON_TER 1 141
SQ SEQUENCE 141 AA; 15785 MW; 4115D3D915751757 CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 141;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
DB 109 NKGTQOYT 116

RESULT 4
ID 085146 PRELIMINARY; PRT; 141 AA.
AC 085146;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.

```

```

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Pöblotzki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
  "XXXSequence variability among different parvovirus B19 isolates.";
  Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; Z70555; CAA94467.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 141
FT NON_TER 1 141
SQ SEQUENCE 141 AA; 15712 MW; 32F6B23B7E6B8583 CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 141;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
DB 106 NKGTQOYT 113

RESULT 5
ID 085168 PRELIMINARY; PRT; 142 AA.
AC 085168;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Pöblotzki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
  "XXXSequence variability among different parvovirus B19 isolates.";
  Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; Z70577; CAA94490.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 142
FT NON_TER 1 142
SQ SEQUENCE 142 AA; 15821 MW; 2B6E4D9A6784F8C4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 142;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYT 8
DB 105 NKGTQOYT 112

RESULT 6
ID 085173 PRELIMINARY; PRT; 142 AA.
AC 085173;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Pöblotzki A., Giegler A., Cassinotti P., Siegl G.,

```

RA WOLF H., Modrow S.;
 RT "XXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70562; CA944495.1;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 142
 SQ SEQUENCE 142 AA; 15820 MW; 3A95B07C0BC0434 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 142;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 105 NKGTOOYT 112

RESULT 7

AC 085138 PRELIMINARY; PRT; 144 AA.
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Gigler A., Caasinocti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70547; CA944459.1;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 144
 SQ SEQUENCE 144 AA; 16022 MW; 95D6EF37BA5997AE CRC64;

Query Match 100.0%; Score 44; DB 12; Length 144;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOOYT 8
 107 NKGTOOYT 114

RESULT 8

ID 085181 PRELIMINARY; PRT; 145 AA.
 AC 085181;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Gigler A., Caasinocti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70550; CA944503.1;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16210 MW; 6B45A1E1B9B23C4 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 109 NKGTOOYT 116

RESULT 9

ID 085161 PRELIMINARY; PRT; 145 AA.
 AC 085161;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Gigler A., Caasinocti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70570; CA944483.1;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16165 MW; B4FE95AB6E4D9A67 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 105 NKGTOOYT 112

RESULT 10

ID 085158 PRELIMINARY; PRT; 146 AA.
 AC 085158;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Gigler A., Caasinocti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70567; CA944480.1;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 146
 SQ SEQUENCE 146 AA; 16252 MW; 2B74FE95AB6E4D9A CRC64;

Query Match 100.0%; Score 44; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
|||||

DB 105 NKGTOOYT 112

RESULT 11

ID Q85142 PRELIMINARY; PRT; 147 AA.

AC Q85142; 01-NOV-1996 (TRENBLREL. 01, Created)

DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)

DE 01-DEC-2001 (TRENBLREL. 19, Last annotation update)

GN VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1] SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,

RT "XXXSequence variability among different parvovirus B19 isolates.";

DL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z70551; CA94463.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 147

FT SEQUENCE 147 AA; 16424 MW; E4ECF2459B308BA9 CRC64;

QY Query Match 100.0%; Score 44; DB 12; Length 147;

Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 108 NKGTOOYT 115

RESULT 12

ID Q85150 PRELIMINARY; PRT; 148 AA.

AC Q85150; 01-NOV-1996 (TRENBLREL. 01, Created)

DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)

DE 01-DEC-2001 (TRENBLREL. 19, Last annotation update)

GN VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1] SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,

RT "XXXSequence variability among different parvovirus B19 isolates.";

DL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z70559; CA94471.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 148

FT SEQUENCE 148 AA; 16539 MW; 8814ECF2459B308B CRC64;

QY Query Match 100.0%; Score 44; DB 12; Length 148;

Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 NKGTOOYT 8
|||||

DB 108 NKGTOOYT 115

RESULT 13

ID Q85131 PRELIMINARY; PRT; 151 AA.

AC Q85131; 01-NOV-1996 (TRENBLREL. 01, Created)

DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)

DE 01-DEC-2001 (TRENBLREL. 19, Last annotation update)

GN VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1] SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,

RT "Sequence variability among different parvovirus B19 isolates.";

DL J. Gen. Virol. 77:1781-1785(1996).

DR EMBL; Z70540; CA94452.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 151

FT SEQUENCE 151 AA; 16902 MW; 3CA74914B8E73A3E CRC64;

QY Query Match 100.0%; Score 44; DB 12; Length 151;

Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 109 NKGTOOYT 116

RESULT 14

ID Q85177 PRELIMINARY; PRT; 151 AA.

AC Q85177; 01-NOV-1996 (TRENBLREL. 01, Created)

DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)

DE 01-DEC-2001 (TRENBLREL. 19, Last annotation update)

GN VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1] SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,

RT "XXXSequence variability among different parvovirus B19 isolates.";

DL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z70586; CA94499.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 151

FT SEQUENCE 151 AA; 16888 MW; 06C6A614A616DDA3 CRC64;

QY Query Match 100.0%; Score 44; DB 12; Length 151;

Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 NKGTOOYT 8
|||||

DB 109 NKGTOOYT 116

RESULT 15

Q85123 PRELIMINARY; PRT; 153 AA.
 ID Q85123;
 AC Q85123;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DE 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96332516; PubMed=8760426;
 RA Hemner A., Von Poblitzki A., Giegler A., Cassinotti P., Siegl G.,
 Wolf H., Modrow S.;
 RT "Sequence variability among different parvovirus B19 isolates";
 RL J. Gen. Virol. 77:1781-1785(1996).
 DR EMBL; Z70532; CAA9444.1; -.
 InterPro; IPR001403; Parvo_coat.
 Pfam; PF00740; Parvo_coat; 1.
 NON_TER 1 153
 FT NON_TER 153
 SQ SEQUENCE 153 AA; 17131 MW; ECCA1F44020814EC CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 44; DB 12; Length 153;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 Db 108 NKGTOOYT 115

RESULT 16
 ID Q85196 PRELIMINARY; PRT; 157 AA.
 AC Q85196;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Hemner A., Von Poblitzki A., Giegler A., Cassinotti P., Siegl G.,
 Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates";
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70603; CAA94518.1; -.
 InterPro; IPR001403; Parvo_coat.
 Pfam; PF00740; Parvo_coat; 1.
 NON_TER 1 157
 FT NON_TER 157
 SQ SEQUENCE 157 AA; 17591 MW; 1D42191887FCE03 CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 44; DB 12; Length 157;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 Db 110 NKGTOOYT 117

RESULT 17
 ID Q85135 PRELIMINARY; PRT; 162 AA.
 AC Q85135;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)

DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemner A., Von Poblitzki A., Giegler A., Cassinotti P., Siegl G.,
 Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates";
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70544; CAA94456.1; -.
 InterPro; IPR001403; Parvo_coat.
 Pfam; PF00740; Parvo_coat; 1.
 NON_TER 1 162
 FT NON_TER 162
 SQ SEQUENCE 162 AA; 18005 MW; 3F51443566660F2 CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 44; DB 12; Length 162;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 Db 109 NKGTOOYT 116

RESULT 18
 ID Q913X0 PRELIMINARY; PRT; 546 AA.
 AC Q913X0;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
 DE Major virus capsid protein VP2 (Fragment).
 GN VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Lat1;
 RA Hokynar K., Soderlund-Venemo M., Ranki A., Kiviluoto O., Partio E.K.,
 Hedman K.;
 RT "A new parvovirus B19 genotype persistent in skin";
 RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY044266; AAK95572.1; -.
 InterPro; IPR001403; Parvo_coat.
 Pfam; PF00740; Parvo_coat; 1.
 NON_TER 546 546
 FT NON_TER 546
 SQ SEQUENCE 546 AA; 59934 MW; 467BE468A57282E5 CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 44; DB 12; Length 546;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 Db 395 NKGTOOYT 402

RESULT 19
 ID Q913W7 PRELIMINARY; PRT; 546 AA.
 AC Q913W7;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
 DE Major virus capsid protein VP2 (Fragment).
 GN VP2.
 OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RN NCBI_TaxID=10798;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HAM;
 RA Hokynar K., Soderlund-Venermo M., Ranki A., Kivlinuoto O., Partio E.K.,
 Hedman K.;
 RT "A new parvovirus B19 genotype persistent in skin."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY044268; AAK5575.1; -;
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON TER 546 546
 SQ SEQUENCE 546 AA; 59910 MW; BBP1DD23695E3FCD CRC64;

Query Match 100.0%; Score 44; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.54;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYT 8
 |||||
 395 NKGTQOYT 402

RESULT 20
 ID Q9PZS9 PRELIMINARY; PRT; 554 AA.
 AC Q9PZS9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE VP2 capsid protein.
 GN VP.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RN NCBI_TaxID=10798;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HV;
 RA Gallinella G., Venturoli S.;
 RT "B19 Genome Sequence and Structure Analysis."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF162273; AAD46615.1; -;
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60833 MW; 7901FEBD65A97B6 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYT 8
 |||||
 395 NKGTQOYT 402

RESULT 21
 ID Q90201 PRELIMINARY; PRT; 554 AA.
 AC Q90201;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE VP2 structural protein (Capsid protein VP2) (Fragment).
 GN VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RN NCBI_TaxID=10798;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HAM;
 RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the

RT 'VP1/VP2 gene from multiple isolates.'
 RL J. Gen. Virol. 77:2767-2774(1996).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MI, and N8;
 RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K.,
 Sasaki T.;
 RT "Sequence of human parvovirus B19 isolates from patients with
 RT rheumatoid arthritis."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N8;
 RA MEDLINE=90218047; PubMed=2157807;
 RX Umene K., Nunoue T.;

RT "The genome type of human parvovirus B19 strains isolated in Japan
 RT during 1981 differs from types detected in 1986 to 1987: a correlation
 RT between genome type and prevalence."
 RL J. Gen. Virol. 71:983-986(1990).

DR EMBL: U53595; AAB47453.1; -;
 DR EMBL: U53596; AAB47455.1; -;
 DR EMBL: U53593; AAB47449.1; -;
 DR EMBL: U53594; AAB47451.1; -;
 DR EMBL: U53597; AAB47457.1; -;
 DR EMBL: U53600; AAB47463.1; -;
 DR EMBL: U53601; AAB47465.1; -;
 DR EMBL: U53598; AAB47459.1; -;
 DR EMBL: U53599; AAB47461.1; -;
 DR EMBL: AB030693; BAA90290.1; -;
 DR EMBL: AB030673; BAA90268.1; -;
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON TER 554 554
 SQ SEQUENCE 554 AA; 60853 MW; 98FE598F20CB6F CRC64;

Query Match 100.0%; Score 44; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYT 8
 |||||
 395 NKGTQOYT 402

RESULT 22
 ID Q65790 PRELIMINARY; PRT; 554 AA.
 AC Q65790;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RN NCBI_TaxID=10798;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VEN1;
 RA Echevarria Mayo J.E., Erdman D.D.;
 RT Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U31358; AAB35559.1; -;
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60798 MW; BA89F2B293BAE24 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYT 8
 |||||
 395 NKGTQOYT 402

```

RESULT 23
Q912B7 PRELIMINARY; PRT; 554 AA.
AC Q912B7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
GN VP2 protein.
OS Human erythrovirus V9.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=72197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V9;
RA Nguyen Q.T.;
RT "Molecular cloning and sequencing of a novel human erythrovirus genome: new species beside B19 in the genus Erythrovirus.";
Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AJ249437; CAC80622.1; -.
InterPro: IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60885 MW; B4F338CDA80F336 CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 554;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
DB 395 NKGTQOY 402

RESULT 24
Q9UGP7 PRELIMINARY; PRT; 554 AA.
AC Q9UGP7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein VP2.
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rm;
RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K., Sasaki T.;
RT "Sequence of human parvovirus B19 isolates from patients with rheumatoid arthritis.";
Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AB030694; BAA90293.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60839 MW; 0DB958B33C73BF64 CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 554;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
DB 395 NKGTQOY 402

RESULT 25
Q9WKL9 PRELIMINARY; PRT; 554 AA.
AC Q9WKL9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)

```

```

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Structural protein VP2.
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20014169; PubMed=10548133;
RA Hemanet A., Beckenlehner K., Wolf H., Lang B., Motrow S.;
RT "Acute parvovirus B19 infection in connection with a flare of systemic lupus erythematoses in a female patient.";
RL J. Clin. Virol. 14:73-77(1999).
DR EMBL; AF113323; AAC99439.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60913 MW; 98FB588FF205C66F CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 554;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
DB 395 NKGTQOY 402

RESULT 26
Q8JN54 PRELIMINARY; PRT; 554 AA.
AC Q8JN54;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D91.1;
RX MEDLINE=22174902; PubMed=12186896;
RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G., Meritet J.F., Garbarg-Chenon A.;
RT "Genetic diversity within Human Erythroviruses: Identification of three genotypes.";
RL J. Virol. 76:9124-9134(2002).
DR EMBL; AY083234; AAL91014.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60830 MW; A298575C083F24C8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 554;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
DB 395 NKGTQOY 402

RESULT 27
Q9PZT8 PRELIMINARY; PRT; 760 AA.
AC Q9PZT8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VP1/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kati12;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT "Integrity and full coding sequence of B19 virus DNA persisting in
RT human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL: AF161224; AAD45912.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 760
SQ SEQUENCE 760 AA; 83403 MW; EA6E0145E3A0E5A CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 760;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 NKGTQOYT 8
|||||
622 NKGTQOYT 629

```

```

RESULT 28
Q9PZU0 PRELIMINARY; PRT; 761 AA.
AC Q9PZU0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VP1/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kati1;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT "Integrity and full coding sequence of B19 virus DNA persisting in
RT human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL: AF161223; AAD45910.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 761
SQ SEQUENCE 761 AA; 83574 MW; 73A84C930D473530 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 761;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 NKGTQOYT 8
|||||
DB 622 NKGTQOYT 629

```

```

RESULT 29
Q9PZT6 PRELIMINARY; PRT; 765 AA.
AC Q9PZT6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-DEC-2001 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VP1/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kati3;

```

```

RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT "Integrity and full coding sequence of B19 virus DNA persisting in
RT human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL: AF161225; AAD45915.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 765
SQ SEQUENCE 765 AA; 83998 MW; 89E2546086DCDBE8 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 765;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 NKGTQOYT 8
|||||
DB 622 NKGTQOYT 629

```

```

RESULT 30
Q9PZT4 PRELIMINARY; PRT; 769 AA.
AC Q9PZT4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VP1/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kati4;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT "Integrity and full coding sequence of B19 virus DNA persisting in
RT human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL: AF161226; AAD45917.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 769
SQ SEQUENCE 769 AA; 84578 MW; 0749D46E5CA7B868 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 769;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 NKGTQOYT 8
|||||
DB 622 NKGTQOYT 629

```

```

RESULT 31
Q913X1 PRELIMINARY; PRT; 773 AA.
AC Q913X1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Minor virus capsid protein VP1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kati;
RA Hokynar K., Soderlund-Venemo M., Ranki A., Kiviluoto O., Partio E.K.,

```


RT "A new parvovirus B19 genotype persistent in skin."
 DR Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 RL EMBL; AY044266; AAK95571.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON TER 773 773
 SQ SEQUENCE 773 AA; 85164 MW; D60CCAGF90B05378 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 44; DB 12; Length 773;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 622 NKGTOOYT 629

RESULT 32
 Q913W8 PRELIMINARY; PRT; 773 AA.

Q913W8; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Minor virus capsid protein VPI (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=HAM;
 RA Hokynar K.; Soderlund-Venermo M., Ranki A., Kivluoto O., Partio E.K., Hedman K.;
 RL "A new parvovirus B19 genotype persistent in skin."
 DR Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY044266; AAK95571.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON TER 773 773
 SQ SEQUENCE 773 AA; 85140 MW; 2886F3245F9CEB50 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 44; DB 12; Length 773;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 622 NKGTOOYT 629

RESULT 33
 Q8JUN56 PRELIMINARY; PRT; 781 AA.

Q8JUN56; 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=D91.1;
 RA MEDLINE=22174902; PubMed=12186896;
 RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G.,
 RA Merlet J.F., Garbary-Chenon A.;
 RT "Genetic Diversity within Human Erythroviruses: Identification of
 RT Three Genotypes";
 RL J. Virol. 76:9124-9134(2002).
 DR EMBL; AY083234; AAL91013.1; -
 DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86142 MW; 4BB71FEDD41FC8F3 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 44; DB 12; Length 781;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 622 NKGTOOYT 629

RESULT 34
 Q65789 PRELIMINARY; PRT; 781 AA.

Q65789; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=VENI;
 RA Echevarria Mayo J.E., Erdman D.D.;
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U31358; AAA83558.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 85985 MW; ELD5FDE230935DD CRC64;

Query Match
 Best Local Similarity 100.0%; Score 44; DB 12; Length 781;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 622 NKGTOOYT 629

RESULT 35
 P90223 PRELIMINARY; PRT; 781 AA.

P90223; 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE VPI and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=USA2;
 RA MEDLINE=97081188; PubMed=8922470;
 RA Erdman D.D., Durigon E.V., Wang Q.Y., Anderson L.V.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the
 RT VPI/VP2 gene from multiple isolates";
 RT J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38514; AAB47796.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86070 MW; CID911E39CF04A88 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 44; DB 12; Length 781;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 622 NKGTOOYT 629

RESULT 36

DB 085191 PRELIMINARY; PRT; 781 AA.
 AC 085191;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE ORF2 protein.
 GN ORF2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemmer A., Von Podolczki A., Gigler A., Casinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates.";
 Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 EMBL; Z70599; CA94513.1; -.
 InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86026 MW; AA02577B683EDB2D CRC64;

Query Match 100.0%; Score 44; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.79; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 622 NKGTOOYT 629

RESULT 37
 P90221 PRELIMINARY; PRT; 781 AA.

AC P90221;
 DT 01-MAY-1997 (TRENBLREL. 03, Created)
 DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)
 DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)
 DE VPI and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 STRAIN=KOR2;
 MEDLINE=97081188; PubMed=8922470;
 Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 "Genetic diversity of human parvovirus B19: sequence analysis of the
 VPI/VP2 gene from multiple isolates.";
 RT VPI/VP2 gene from multiple isolates.";
 RL J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38511; AAB47791.1; -.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86070 MW; 9FA830083FEF1357 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.79; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 622 NKGTOOYT 629

RESULT 38
 P89318 PRELIMINARY; PRT; 781 AA.
 AC P89318;
 DT 01-MAY-1997 (TRENBLREL. 03, Created)
 DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)

DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)
 DE VPI and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAP1;
 RX MEDLINE=97081188; PubMed=8922470;
 Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 "Genetic diversity of human parvovirus B19: sequence analysis of the
 VPI/VP2 gene from multiple isolates.";
 RT VPI/VP2 gene from multiple isolates.";
 RL J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38509; AAB47791.1; -.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86064 MW; 3AB65CE6906339 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.79; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 622 NKGTOOYT 629

RESULT 39
 O8JYE3 PRELIMINARY; PRT; 781 AA.
 AC O8JYE3;
 DT 01-OCT-2002 (TRENBLREL. 22, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE 7.5 kDa protein.
 OS Erythrovirus A6.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_TaxID=182494;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen Q.T., Wong S., Brown K.E.;
 RT "Identification and characterization of a second novel human
 RT erythrovirus variant, A6.";
 RL Virology 0:0-0(2002).
 DR EMBL; AY064475; AAL55418.1; -.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86277 MW; E73AB663E4551128 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.79; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYT 8
 DB 622 NKGTOOYT 629

RESULT 40
 P90224 PRELIMINARY; PRT; 781 AA.
 AC P90224;
 DT 01-MAY-1997 (TRENBLREL. 03, Created)
 DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)
 DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)
 DE VPI and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BR21;
 RX MEDLINE=97081188; PubMed=8922470;

RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.U.;
RT "Genetic diversity of human parvovirus BL9: sequence analysis of the
VP1/VP2 gene from multiple isolates."
RL J. Gen. Virol. 77:2767-2774(1996).
DR EMBL; U38546; AAB47801.1; -;
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86039 MW; C3A29BE1DB8AD378 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.79; 0; Mismatches 0; Gaps 0;
Matches 8; Conservative 0; Indels 0;

Oy 1 NKGTQOYT 8
|||
Db 622 NKGTQOYT 629

Search completed: August 20, 2003, 09:23:59
Time : 23.2577 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 20.4124 Seconds
(without alignments)
46.656 Million cell updates/sec

Title: US-09-991-433-7

Perfect score: 33

Sequence: 1 QOYTDQ 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	370	12	AA13406
2	33	100.0	433	23	ABP73733
3	33	100.0	543	12	AA13405
4	33	100.0	554	16	AAW08987
5	33	100.0	554	20	AAV23230
6	33	100.0	554	24	ABP57264
7	33	100.0	554	24	ABP57267
8	33	100.0	756	21	AAV71231
9	33	100.0	781	16	AAW08986

10	33	100.0	781	20	AAV23227
11	33	100.0	781	24	ABP57263
12	33	100.0	781	24	ABP57266
13	29	87.9	106	23	ABP60721
14	29	87.9	271	23	ABBS4456
15	29	87.9	414	23	ABBS3614
16	29	87.9	545	23	ABP64705
17	29	87.9	545	24	ABU54560
18	29	87.9	545	24	ABU54561
19	29	87.9	545	24	ABJ76653
20	29	87.9	545	24	ABU12117
21	29	87.9	640	20	AAV39926
22	29	87.9	1778	22	ABBS2677
23	28	84.8	91	23	ABP69513
24	28	84.8	100	21	AAW02653
25	28	84.8	119	21	AA11198
26	28	84.8	119	23	ABP08647
27	28	84.8	120	21	AA10236
28	28	84.8	130	21	AA10235
29	28	84.8	169	23	ABBS4752
30	28	84.8	175	21	AA10234
31	28	84.8	196	18	AAW28253
32	28	84.8	259	23	ABG60924
33	28	84.8	261	23	AAU82959
34	28	84.8	312	22	AAU34323
35	28	84.8	312	23	ABP40664
36	28	84.8	325	22	AAU33364
37	28	84.8	325	22	AAU33354
38	28	84.8	384	23	ABBS1551
39	28	84.8	447	22	ABBS5406
40	28	84.8	493	20	AAV22206
41	28	84.8	572	13	AA26574
42	28	84.8	651	11	AA106460
43	28	84.8	651	14	AA137359
44	28	84.8	651	17	AAW06419
45	28	84.8	651	20	AAV23197

ALIGNMENTS

RESULT 1
ID AAR13406 standard; Protein; 370 AA.
AC AAR13406;
DT 24-OCT-1991 (first entry)
DE Parvo virus B19 PANSE.
KW Primer; PCR; PAPST; globulin.
OS Synthetic.
PN DE4003826-A.
PD 14-AUG-1991.
PP 08-FEB-1990; 90DE-4003826.
PR 08-FEB-1990; 90DE-4003826.
PA (MIKRO) MIKROGEN MOLEKULARB.
Soutschek E, Motz M;
WPI; 1991-246423/34.
Immunologically active parvo virus B19 peptide(s) - comprising
Capid protein VPI or VP2 fragments, useful for antibody
detection or vaccination

Erythrovirus V9 VP
Human parvovirus B
Human parvovirus B
Anacystis nidulans
Lactococcus lactis
Herbicideally activ
Human NOXV polyep
Human NOXV polyep
Human protein modi
Human protein modi
Bovine rotavirus S
Escherichia coli P
Human polypeptide
Human secreted pro
Human ORFX ORF62
Human ORFX protein
Arabidopsis thalia
Arabidopsis thalia
Lactococcus lactis
Arabidopsis thalia
Amino acid sequenc
Novel floral meris
Human homologue of
Staphylococcus aur
Staphylococcus epi
Staphylococcus aur
Herbicidally activ
Drosophila melanog
Amino acid sequenc
Sequence encoded b
BPGS1208 protoxin
B isolate 43F. B
Antiscarb pest to
Amino acid sequenc

PS Claim 16; Page 10-11; 22pp; German.
 XX
 CC VP2 (AAR13405) and its fragments PANSE and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AA013159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 XX
 SQ Sequence 370 AA;
 QY Query Match 100.0%; Score 33; DB 12; Length 370;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 QOYTDQ 6
 235 QOYTDQ 240
 RESULT 2
 ID ABP73733 standard; Protein; 433 AA.
 XX ABP73733;
 XX 30-JUN-2003 (first entry)
 XX
 DE Candida albicans essential protein SEQ ID NO 7570.
 XX
 XX Fungus; Yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
 KM signal transduction; DNA replication; cell division; growth;
 KM proliferation; Candida albicans; fungicide; antifungal.
 XX
 OS Candida albicans.
 XX
 PN MO200253728-A2.
 XX
 PD 11-JUL-2002.
 XX
 PF 26-DEC-2001; 2001WO-US49486.
 XX
 PR 29-DEC-2000; 2000US-259128P.
 XX
 PR 20-FEB-2001; 2001US-0792024.
 XX
 PR 22-AUG-2001; 2001US-314050P.
 XX
 (ELIT-) ELITRA PHARM INC.
 XX
 PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
 XX
 DR WPI; 2002-566694/60.
 DR
 DR N-PSDB; AB232283.
 XX
 PT Constructing strains for identifying gene products as effective targets
 PT for therapeutic intervention, by inactivating in the strain one allele
 PT of a gene and placing other allele of the gene under conditional
 PT expression -
 XX
 PS Claim 44; SEQ ID NO 7570; 167pp + Sequence listing; English.
 XX
 CC The invention relates to constructing (M1) a strain of diploid fungal
 CC cells in which both alleles of a gene are modified, comprising modifying
 CC one allele by insertion or replacement by a cassette having an
 CC expressible selectable marker and modifying other allele by
 CC recombination, of a promoter replacement fragment with a heterologous
 CC promoter, so that expression of the second allele is regulated by the
 CC promoter. (M1) is useful for constructing a strain of diploid fungal
 CC cells in which both alleles of a gene are modified. The diploid fungal
 CC cells having both alleles modified are useful for identifying a gene that
 CC is essential to the survival or growth of a fungus, a gene that

CC contributes to the virulence and/or pathogenicity of a fungus, a gene
 CC that contributes to the resistance of a diploid fungus to an antifungal
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus
 CC and for identifying a therapeutic agent for treatment of a mammalian
 CC disease. (M1) is useful for identifying a compound which modulates the
 CC activity of a gene product, preferably enzymatic activity, carbon
 CC compound catabolism, biosynthetic, transporter, transcriptional,
 CC translational, signal transduction, DNA replication and cell division
 CC activity. The method is useful for identifying a compound having the
 CC ability to inhibit growth or proliferation of C. albicans cells and for
 CC treating infection by C. albicans. The present sequence is that of an
 CC essential Candida albicans protein used in the method of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office.
 XX
 SQ Sequence 433 AA;
 QY Query Match 100.0%; Score 33; DB 23; Length 433;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 QOYTDQ 6
 391 QOYTDQ 396
 RESULT 3
 ID AAR13405 standard; Protein; 543 AA.
 XX AAR13405;
 XX
 AC AAR13405;
 XX
 DT 24-OCT-1991 (first entry)
 XX
 DE Parvo virus B19 VP2.
 XX
 XX
 KM Primer; PCR; globulin; PANSE; PAPST.
 XX
 OS Synthetic.
 XX
 PN DE4003826-A.
 XX
 PD 14-AUG-1991.
 XX
 PF 08-FEB-1990; 90DE-4003826.
 XX
 PR 08-FEB-1990; 90DE-4003826.
 XX
 PA (MIR-) MIKROGEN MOLEKULARB.
 XX
 PI Soutschek E, Motz M;
 XX
 DR WPI; 1991-246423/34.
 DR
 DR
 PT Immunologically active parvo virus B19 peptide(s) - comprising
 PT capsid protein VP1 or VP2 fragments, useful for antibody
 PT detection or vaccination
 XX
 PS Disclosure; Fig 2-6; 22pp; German.
 XX
 CC VP2 and its fragments PANSE (AAR13406) and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AA013159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 XX
 SQ Sequence 543 AA;

Query Match 100.0%; Score 33; DB 12; Length 543;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 |||||
 DB 408 QOYTDQ 413

RESULT 4

AAW08987

ID AAW08987 standard; Protein; 554 AA.

XX AAW08987;

DT 27-FEB-1997 (first entry)

DE Human parvovirus VP-2 protein.

XX

Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 erythblastemia; abortion; universal fetal hydrops; liver disease;
 haemorrhagic fever; rheumatism; detection; IgG antibody.

KM Human parvovirus.

OS JP07147986-A.

PN 13-JUN-1995.

XX 24-SEP-1992; 92JP-0281017.

XX 24-SEP-1992; 92JP-0281017.

PR (DENK-) DENKA SEIKEN KK.

PA (ELSD) DENKI KAGAKU KOGYO KK.

DR WPI; 1995-242756/32.

XX N-PSDB; AAT49535.

Human parvovirus gene coding for a polypeptide - useful for
 developing vaccines against parvoviral diseases such as
 erythroblastemia, haemorrhagic fever, etc.

Claim 3; Page 7-9; 38pp; English.

The sequences given in AAW08986 represent the parvovirus structural
 proteins, VP-1 and VP2, and the non-structural protein, NS. The
 genomic fragment encoding these proteins may be used for the
 diagnosis and development of vaccines for parvoviral diseases including
 erythblastemia, abortion, universal fetal hydrops, liver diseases,
 haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 proteins may be used to detect parvovirus IgG antibodies.

SQ Sequence 554 AA;

Query Match 100.0%; Score 33; DB 16; Length 554;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 |||||
 DB 399 QOYTDQ 404

RESULT 5

AAW23230

ID AAW23230 standard; Protein; 554 AA.

XX AAW23230;

DT 26-AUG-1999 (first entry)

XX

DE Erythrovirus V9 VP2 protein.

XX Erythrovirus V9; differential diagnosis; parvovirus; infection;

KM erythrovirus screening; typing; immunoassay; VP2 protein.

OS Erythrovirus.

PN FR2771751-A1.

XX 04-JUN-1999.

XX 03-DEC-1997; 97FR-0015197.

XX 03-DEC-1997; 97FR-0015197.

PA (ASSI-) ASSISTANCE PUBLIQUE HOPITALUX PARIS.

PI Auguste V, Garbarg CA, Nguyen QT;

DR WPI; 1999-349543/30.

XX N-PSDB; AAX81586.

Erythrovirus V9 and its nucleic acid sequences - can be used in the
 diagnosis of its infections

Claim 19; Page 57-58; 80pp; French.

The present sequence represents an erythrovirus V9 protein.
 Probes and primers derived from erythrovirus V9 polynucleotide
 sequences (AAX81580) can be used for differential diagnosis of
 erythrovirus (parvovirus) infections by a combination of
 amplification and hybridisation assay. The probes can also be
 used to assess susceptibility to erythrovirus infection and
 for erythrovirus screening and typing. The antibodies can be
 used in immunoassays for diagnosis of erythrovirus V9 infections.

SQ Sequence 554 AA;

Query Match 100.0%; Score 33; DB 20; Length 554;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 |||||
 DB 399 QOYTDQ 404

RESULT 6

ABP57264

ID ABP57264 standard; Protein; 554 AA.

XX ABP57264;

DT 22-APR-2003 (first entry)

DE Human parvovirus B19 clone B1-VP2 amino acid sequence SEQ ID NO:29.

KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.

OS Human parvovirus B19.

PN WO2003002753-A2.

XX 09-JAN-2003.

XX 28-JUN-2002; 2002WO-US20684.

XX 28-JUN-2001; 2001US-302077P.

PR 19-MAR-2002; 2002US-365956P.

XX 29-MAR-2002; 2002US-369224P.

PA (CHIR) CHIRON CORP.

PI Pichuanes S, Shyamala V;
 XX
 DR WPI; 2003-201510/19.
 DR N-PSDB; ABZ59574.
 XX
 PT Detecting a human parvovirus B19 infection in a biological sample to
 PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -
 XX
 PS Example 4; Fig 7B; 148bp; English.
 XX
 CC The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 554 AA;
 Query Match 100.0%; Score 33; DB 24; Length 554;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 |||||
 Db 399 QOYTDQ 404
 RESULT 7
 ABP57267
 ABP57267 standard; Protein; 554 AA.
 ABP57267;
 DT 22-APR-2003 (first entry)
 XX
 DE Human parvovirus B19 clone B6-VP2 amino acid sequence SEQ ID NO:35.
 XX
 KW Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
 XX
 OS Human parvovirus B19.
 XX
 PN WO2003002753-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WO-US20684.
 XX
 PR 28-JUN-2001; 2001US-302077P.
 PR 19-MAR-2002; 2002US-365956P.
 PR 29-MAR-2002; 2002US-369224P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Pichuanes S, Shyamala V;
 XX

DR WPI; 2003-201510/19.
 DR N-PSDB; ABZ59577.
 XX
 PT Detecting a human parvovirus B19 infection in a biological sample to
 PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -
 XX
 PS Example 4; Fig 10B; 148bp; English.
 XX
 CC The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 554 AA;
 Query Match 100.0%; Score 33; DB 24; Length 554;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 |||||
 Db 399 QOYTDQ 404
 RESULT 8
 AAY71231
 ID AAY71231 standard; Protein; 756 AA.
 AC AAY71231;
 XX
 DT 08-SEP-2000 (first entry)
 XX
 DE Capsid protein encoded by AAV2/human parvovirus B19 chimeric vector.
 XX
 KW Adeno-associated virus; AAV2; human parvovirus B19; chimeric;
 KW recombinant parvoviral vector; cellular tropism; cap protein;
 KW capsid; gene delivery; gene therapy; VP1; VP2; VP3.
 XX
 OS Chimeric - Adeno associated virus serotype 2.
 XX
 PN WO200028004-A1.
 XX
 PD 18-MAY-2000.
 XX
 PF 10-NOV-1999; 99WO-US26505.
 XX
 PR 10-NOV-1998; 98US-0107840.
 PR 10-MAR-1999; 99US-0123651.
 XX
 PA (UNIC-) UNIV NORTH CAROLINA.
 XX
 PI Rabinowitz JE, Samulski RJ, Xiao W;
 XX

DR WPI; 2000-376523/32.
 DR N-PSDB; AAD00833.
 XX Recombinant parvoviral vectors with altered packaging, tropisms and
 PT immunogenic properties, useful in gene therapy protocols -
 XX
 PS Example 21; Page 142; 153pp; English.
 CC The patent discloses modified parvovirus vectors with advantageous
 CC antigenic properties, packaging capabilities and cellular tropisms.
 CC These vectors can be used in standard recombinant DNA protocols e.g. gene
 CC therapy for delivering nucleic acids to cells.
 CC The present sequence is a capsid protein encoded by an
 CC adeno-associated virus serotype 2 (AAV2)/human parvovirus B19 chimeric
 CC vector. This vector encodes AAV2 VP1 and VP2 capsid proteins
 CC and human parvovirus B19 VP2 protein. The chimeric vector was
 CC constructed by replacing the VP3 major cap protein of AAV2 with B19's
 CC VP2. Recombinant parvovirus comprising the chimeric capsid is useful
 CC for gene delivery.

Sequence 756 AA;

Query Match 100.0%; Score 33; DB 21; Length 756;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 |||||
 DB 601 QOYTDQ 606

RESULT 9

AAW08986
 ID AAW08986 standard; Protein; 781 AA.

AC AAW08986;

DT 27-FEB-1997 (first entry)

DE Human parvovirus VP-1 protein.

KW Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 KW non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 KW erythblastemia; abortion; universal fetal hydrops; liver disease;
 KW haemorrhagic fever; rheumatism; detection; IgG antibody.

OS Human parvovirus.

Key Location/Qualifiers

Misc-difference 19
 /label= GLY, Val

FT Misc-difference 61
 /label= Asn, Asp

FT Misc-difference 220
 /label= His, Asn

PN JP07147986-A.

PD 13-JUN-1995.

PF 24-SEP-1992; 92JP-0281017.

PR 24-SEP-1992; 92JP-0281017.

PA (DENK-) DENKA SEIKEN KK.

PA (LEBD) DENKI KAGAKU KOGYO KK.

DR WPI; 1995-242756/32.

DR N-PSDB; AAT49535.

PT Human parvovirus gene coding for a polypeptide - useful for
 PT developing vaccines against parvoviral diseases such as
 PT erythroblastemia, haemorrhagic fever, etc.

XX Claim 2; Page 5-7; 38pp; English.

PS The sequences given in AAW08986 represent the parvovirus structural
 CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.

SQ Sequence 781 AA;

Query Match 100.0%; Score 33; DB 16; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 |||||
 DB 626 QOYTDQ 631

RESULT 10

AAW23227
 ID AAW23227 standard; Protein; 781 AA.

AC AAW23227;

DT 26-AUG-1999 (first entry)

DE Erythrovirus V9 VP1 protein.

KW Erythrovirus V9; differential diagnosis; parvovirus; infection;
 KW erythrovirus screening; typing; immunoassay; VP1 protein.

OS Erythrovirus.

PN FR2771751-A1.

PD 04-JUN-1999.

PF 03-DEC-1997; 97FR-0015197.

PR 03-DEC-1997; 97FR-0015197.

PA (ASST-) ASSISTANCE PUBLIQUE HOPITALUX PARIS.

PI Auguste V, Garbarg CA, Nguyen QT;

DR WPI; 1999-349543/30.

DR N-PSDB; AAX81583.

PT Erythrovirus V9 and its nucleic acid sequences - can be used in the
 PT diagnosis of its infections

PS Claim 19; Page 50-52; 80pp; French.

CC The present sequence represents an erythrovirus V9 protein.
 CC Probes and primers derived from erythrovirus V9 polynucleotide
 CC sequences (AAX81580) can be used for differential diagnosis of
 CC erythrovirus (parvovirus) infections by a combination of
 CC amplification and hybridisation assay. The probes can also be
 CC used to assess susceptibility to erythrovirus infection and
 CC for erythrovirus screening and typing. The antibodies can be
 CC used in immunoassays for diagnosis of erythrovirus V9 infections.

SQ Sequence 781 AA;

Query Match 100.0%; Score 33; DB 20; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6

Db 626 QOYTDQ 631

RESULT 11

ABP57263 standard; Protein; 781 AA.

AC ABP57263;

DT 22-APR-2003 (first entry)

DE Human parvovirus B19 clone B1-VP1 amino acid sequence SEQ ID NO:27.

DE Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.

OS Human parvovirus B19.

PN WO2003002753-A2.

09-JAN-2003.

28-JUN-2002; 2002MO-US20684.

28-JUN-2001; 2001US-302077P.

19-MAR-2002; 2002US-365956P.

29-MAR-2002; 2002US-369224P.

(CHIR) CHIRON CORP.

Pichuanes S, Shyamala V;

WPI; 2003-201510/19.

DR N-PSDB; ABZ59573.

PT Detecting a human parvovirus B19 infection in a biological sample to

PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic

PT acid with a primer complementary to the 3'-terminal portion of the RNA

PT target sequence -

PS Example 4; Fig 6B; 148bp; English.

XX The present invention describes a method for detecting a human parvovirus

XX B19 infection in a biological sample. The method comprises reacting the

XX isolated parvovirus B19 nucleic acid with a first oligonucleotide

XX consisting of a first primer containing a complexing sequence

XX sufficiently complementary to the 3'-terminal portion of the RNA target

XX sequence to complex with. Also described: (1) amplifying a target

XX parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one

XX of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to

XX ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair

XX sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer

XX consisting of a promoter region recognised by a DNA-dependent RNA

XX polymerase operably linked to a human parvovirus B19-specific complexing

XX sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a

XX parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked

XX to an acridinium ester label; and (6) a diagnostic test kit comprising an

XX oligonucleotide primer of (4), and instructions for conducting the

XX diagnostic test. The method is useful for detecting parvovirus infection

XX in a biological sample, such as in blood products, to prevent

XX transmission of the virus through blood and plasma derivatives or by

XX close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267

XX represent sequences used in the exemplification of the present invention.

XX Sequence 781 AA;

SO Query Match 100.0%; Score 33; DB 24; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOYTDQ 6

Db 626 QOYTDQ 631

RESULT 12

ABP57266 standard; Protein; 781 AA.

AC ABP57266;

DT 22-APR-2003 (first entry)

DE Human parvovirus B19 clone B6-VP1 amino acid sequence SEQ ID NO:33.

DE Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.

OS Human parvovirus B19.

PN WO2003002753-A2.

09-JAN-2003.

28-JUN-2002; 2002MO-US20684.

28-JUN-2001; 2001US-302077P.

19-MAR-2002; 2002US-365956P.

29-MAR-2002; 2002US-369224P.

(CHIR) CHIRON CORP.

Pichuanes S, Shyamala V;

WPI; 2003-201510/19.

DR N-PSDB; ABZ59576.

PT Detecting a human parvovirus B19 infection in a biological sample to

PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic

PT acid with a primer complementary to the 3'-terminal portion of the RNA

PT target sequence -

PS Example 4; Fig 9B; 148bp; English.

XX The present invention describes a method for detecting a human parvovirus

XX B19 infection in a biological sample. The method comprises reacting the

XX isolated parvovirus B19 nucleic acid with a first oligonucleotide

XX consisting of a first primer containing a complexing sequence

XX sufficiently complementary to the 3'-terminal portion of the RNA target

XX sequence to complex with. Also described: (1) amplifying a target

XX parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one

XX of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to

XX ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair

XX sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer

XX consisting of a promoter region recognised by a DNA-dependent RNA

XX polymerase operably linked to a human parvovirus B19-specific complexing

XX sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a

XX parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked

XX to an acridinium ester label; and (6) a diagnostic test kit comprising an

XX oligonucleotide primer of (4), and instructions for conducting the

XX diagnostic test. The method is useful for detecting parvovirus infection

XX in a biological sample, such as in blood products, to prevent

XX transmission of the virus through blood and plasma derivatives or by

XX close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267

XX represent sequences used in the exemplification of the present invention.

XX Sequence 781 AA;

SO Query Match 100.0%; Score 33; DB 24; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOYTDQ 6

Db 626 QOYTDQ 631

PN WO200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 XX 28-AUG-2001; 2001WO-EP09992.
 XX
 PR 28-AUG-2001; 2001WO-EP09992.
 XX
 PA (PARB) BAYER AG.
 XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX

Claim 5; SEQ ID NO 2825; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins
 CC (AB90790-AB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 CC

CC Sequence 414 AA;

Query Match 87.9%; Score 29; DB 23; Length 414;
 Best Local Similarity 83.3%; Pred. No. 5.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOYTDQ 6
 XX |||||:
 DB 280 QOYTDK 285

RESULT 16
 ABP64705
 ID ABP64705 standard; Protein; 545 AA.
 XX
 AC ABP64705;

25-FEB-2003 (first entry)

Human protein SEQ ID 365.

XX Human; expressed sequence tag; EST;
 KW haemotopoietic disorder; central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; necrotic;
 KW anti-allergic; anti-inflammatory; immunosuppressive; neuroprotective;
 KW cytoskeletal; haemostatic; vitruclide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.
 XX

OS Homo sapiens.

XX WO200259260-A2.

XX 01-AUG-2002.

XX 16-NOV-2001; 2001WO-US42950.

XX 17-NOV-2000; 2000US-0714936.

XX (HYSE-) HYSEQ INC.

PI Tang YF, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Dermanac RT;
 XX
 DR WPI; 2002-590824/63.
 XX
 DR N-PSDB; ABQ9291.
 XX
 PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity -
 XX
 PS Claim 20; SEQ ID 365; 394pp; English.

CC The present invention relates to novel human coding sequences
 CC (ABQ9268-ABQ99608) and proteins (ABP64682-ABP65022). The sequences are
 CC useful in therapeutic, diagnostic and research method. The
 CC polynucleotides may be used in the field of molecular biology as
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
 CC for the recombinant production of protein, or in generation of anti-sense
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
 CC sequence tags (ESTs) for identifying expressed genes or for physical
 CC mapping of the human genome. The proteins may be used as molecular weight
 CC markers, or as nutritional sources or supplements. The proteins may be
 CC used to maintain and expand cell population in a totipotent or
 CC pluripotent state useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
 CC development of bio-sensors. The polynucleotides and proteins are useful
 CC for preventing, treating or ameliorating disorders involving aberrant
 CC protein expression or biological activity, e.g. hematopoietic disorders,
 CC central/peripheral nervous system diseases, mechanical and traumatic
 CC disorders, non-healing wounds, immune deficiencies and fungal infection,
 CC infectious diseases caused by viral, bacterial and fungal infection,
 CC autoimmune disorders, allergic reactions and conditions, coagulation
 CC disorders, or cancer. The polynucleotide sequences of the invention were
 CC assembled from ESTs isolated mainly by sequencing by hybridization, and
 CC in some cases, sequences obtained from one or more public databases.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPD
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 545 AA;

Query Match 87.9%; Score 29; DB 23; Length 545;
 Best Local Similarity 83.3%; Pred. No. 7.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOYTDQ 6
 XX |||||:
 DB 411 QOYTDK 416

RESULT 17
 ABUS4560
 ID ABUS4560 standard; Protein; 545 AA.
 XX
 AC ABUS4560;

03-JUN-2003 (first entry)

Human NOVX polypeptide #19.

XX Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
 KW hypertension; congenital heart defect; aortic stenosis; valve disease;
 KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
 KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
 KW tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;
 KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
 KW Parkinson's disease; immune disorder; haemotopoietic disorder;
 KW haemophilia; hypercoagulation; Crohn's disease; cancer.
 XX

OS Homo sapiens.

XX WO200281498-A2.

PD 17-OCT-2002.
 XX 03-APR-2002; 2002WO-US10780.
 XX
 PR 03-APR-2001; 2001US-281086P.
 PR 03-APR-2001; 2001US-281136P.
 PR 05-APR-2001; 2001US-28163P.
 PR 05-APR-2001; 2001US-281906P.
 PR 06-APR-2001; 2001US-282020P.
 PR 10-APR-2001; 2001US-282930P.
 PR 10-APR-2001; 2001US-282934P.
 PR 12-APR-2001; 2001US-283512P.
 PR 13-APR-2001; 2001US-283710P.
 PR 17-APR-2001; 2001US-284234P.
 PR 19-APR-2001; 2001US-285325P.
 PR 20-APR-2001; 2001US-285381P.
 PR 20-APR-2001; 2001US-285609P.
 PR 23-APR-2001; 2001US-285748P.
 PR 23-APR-2001; 2001US-285890P.
 PR 24-APR-2001; 2001US-286068P.
 PR 25-APR-2001; 2001US-286292P.
 PR 27-APR-2001; 2001US-287213P.
 PR 02-MAY-2001; 2001US-288257P.
 PR 29-MAY-2001; 2001US-294164P.
 PR 30-MAY-2001; 2001US-294484P.
 PR 18-JUN-2001; 2001US-298952P.
 PR 19-JUN-2001; 2001US-299237P.
 PR 19-JUN-2001; 2001US-299276P.
 PR 12-SEP-2001; 2001US-318750P.
 PR 25-SEP-2001; 2001US-324800P.
 PR 25-SEP-2001; 2001US-324802P.
 PR 27-SEP-2001; 2001US-325684P.
 PR 17-OCT-2001; 2001US-330143P.
 PR 14-NOV-2001; 2001US-332131P.
 PR 14-NOV-2001; 2001US-332240P.
 PR 14-NOV-2001; 2001US-332779P.
 PR 21-NOV-2001; 2001US-332115P.
 PR 04-DEC-2001; 2001US-337621P.
 PR 03-JAN-2002; 2002US-345783P.
 PR 16-JAN-2002; 2002US-350251P.
 PR 02-APR-2002; 2002US-0114270.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA, Patturajan M, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V, Padigaru M, Shimkete RA, Gangoli EA, Taupier RJ, Casman SJ, Ji W, Anderson DM, Lette MW, Rastelli L, Edinger SR, Stone DJ, MacDougall JR, Rothenberg ME, Mazur A, Miller I, Peyman JA, Ellerman K;
 XX
 XX WPI; 2003-046858/04.
 DR N-PSDB; ABX72188.
 XX
 PT New isolated NOVX polypeptide useful for treating atherosclerosis, metabolic disorders, diabetes, obesity, infectious diseases, anorexia, neurodegenerative disorders, Alzheimer's disease and cancer
 PT neurodegenerative disorders, Alzheimer's disease and cancer
 PS Claim 1; Page 127; 666pp; English.
 XX
 CC The invention relates to human polypeptides, termed NOVX, and the CC polynucleotides encoding them. The polypeptides and polynucleotides are useful for diagnosing disease, and screening for potential therapeutic agents. The sequences are useful for treating metabolic disorders, CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic stenosis, atrial septal defect (ASD), atrioventricular canal defect, CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma, CC atherosclerosis, obesity, infectious diseases, anorexia, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders, CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease and cancer. Sequences ABUS4542-ABUS4647 represent human NOVX polypeptides

CC of the invention.
 XX
 SQ Sequence 545 AA;
 Query Match 87.9%; Score 29; DB 24; Length 545;
 Best Local Similarity 83.3%; Pred. No. 7.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 DB 411 QOYTDQ 416
 RESULT 18
 ABUS4561
 ID ABUS4561 standard; Protein; 545 AA.
 XX
 AC ABUS4561;
 XX
 DT 03-JUN-2003 (first entry)
 XX
 DE Human NOVX polypeptide #20.
 XX
 KW Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
 KW hypertension; congenital heart defect; aortic stenosis; valve disease;
 KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
 KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
 KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
 KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
 KW Parkinson's disease; immune disorder; haematopoietic disorder;
 KW haemophilia; hypercoagulation; Crohn's disease; cancer.
 OS Homo sapiens.
 XX
 FN WO200281498-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 03-APR-2002; 2002WO-US10780.
 XX
 PR 03-APR-2001; 2001US-281086P.
 PR 03-APR-2001; 2001US-281136P.
 PR 05-APR-2001; 2001US-28163P.
 PR 05-APR-2001; 2001US-281906P.
 PR 06-APR-2001; 2001US-282020P.
 PR 10-APR-2001; 2001US-282930P.
 PR 10-APR-2001; 2001US-282934P.
 PR 12-APR-2001; 2001US-283512P.
 PR 13-APR-2001; 2001US-283710P.
 PR 17-APR-2001; 2001US-284234P.
 PR 19-APR-2001; 2001US-285325P.
 PR 20-APR-2001; 2001US-285381P.
 PR 23-APR-2001; 2001US-285748P.
 PR 23-APR-2001; 2001US-285890P.
 PR 24-APR-2001; 2001US-286068P.
 PR 25-APR-2001; 2001US-286292P.
 PR 27-APR-2001; 2001US-287213P.
 PR 02-MAY-2001; 2001US-288257P.
 PR 29-MAY-2001; 2001US-294164P.
 PR 30-MAY-2001; 2001US-294484P.
 PR 18-JUN-2001; 2001US-298952P.
 PR 19-JUN-2001; 2001US-299237P.
 PR 19-JUN-2001; 2001US-299276P.
 PR 12-SEP-2001; 2001US-318750P.
 PR 25-SEP-2001; 2001US-324802P.
 PR 25-SEP-2001; 2001US-324802P.
 PR 27-SEP-2001; 2001US-325684P.
 PR 17-OCT-2001; 2001US-330143P.
 PR 14-NOV-2001; 2001US-332131P.
 PR 14-NOV-2001; 2001US-332240P.
 PR 14-NOV-2001; 2001US-332779P.
 PR 21-NOV-2001; 2001US-332115P.

PR 04-DEC-2001; 2001US-337621P.
 PR 03-JAN-2002; 2002US-345783P.
 PR 16-JAN-2002; 2002US-350251P.
 PR 02-APR-2002; 2002US-0114270.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Guo X, Kekuda R, Miller CE, Malynkar UM, Spytek KA, Patirajan M;
 PI Liu X, Guev VY, Li L, Verner CM, Zehusen BD, Gorman L;
 PI Shenoy SG, Pena CE, Smithson G, Burgess CB, Gerlach V;
 PI Padigaru M, Shinkets RA, Gangoli EA, Taupier RJ, Casman SJ, Ji W;
 PI Anderson DM, Lette MW, Rastelli L, Edinger SR, Stone DJ;
 PI MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
 PI Ellerman K;
 XX
 DR WPI; 2003-046858/04.
 DR N-PSDB; ABX72189.
 XX
 PT New isolated NOVX polypeptide useful for treating atherosclerosis,
 PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
 PT neurodegenerative disorders, Alzheimer's disease and cancer -
 XX
 PS Claim 1; Page 128; 666pp; English.
 XX
 CC The invention relates to human polypeptides, termed NOVX, and the
 CC polynucleotides encoding them. The polypeptides and polynucleotides are
 CC useful for diagnosing disease, and screening for potential therapeutic
 CC agents. The sequences are useful for treating metabolic disorders,
 CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
 CC stenosis, atrial septal defect (ASD), atriocentricular canal defect,
 CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
 CC septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,
 CC atherosclerosis, obesity, infectious diseases, anorexia, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
 CC haemotopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
 CC and cancer. Sequences ABU54542-ABU54647 represent human NOVX polypeptides
 CC of the invention.
 CC
 SO Sequence 545 AA;
 XX
 QY Query Match 87.9%; Score 29; DB 24; Length 545;
 DB Best Local Similarity 83.3%; Pred. No. 7.5e+02;
 DB 411 QOYTD 416 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 OLIT 19
 ABJ26653
 ID ABJ26653 standard; Protein; 545 AA.
 XX
 AC ABJ26653;
 XX
 DT 01-MAY-2003 (first entry)
 DE Human protein modification + maintenance molecule protein SEQ ID No 7.
 XX
 KW Cytostatic; antiarteriosclerotic; anti-HIV; antiallergic; anticonvulsant;
 KW cerebroprotective; antiparkinsonian; nootropic; antiinflammatory;
 KW anticancer; hepatotropic; gynaecological; antibacterial; vitruide;
 KW protozoicide; antiparasitic; cell proliferative disease; PMOD;
 KW protein modification and maintenance molecule; immunogenic fragment;
 KW cancer; autoimmune; inflammatory disease; neurological disorder;
 KW gastrointestinal; developmental; vesicle trafficking disorder; infection;
 KW protein-protein interaction; drug-target interaction;
 KW gene expression profile; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003000844-A2.
 XX

PD 03-JAN-2003.
 XX
 XX 18-JUN-2002; 2002WO-US19360.
 XX
 PR 22-JUN-2001; 2001US-300508P.
 PR 06-JUL-2001; 2001US-303445P.
 PR 13-JUL-2001; 2001US-305405P.
 PR 09-AUG-2001; 2001US-311442P.
 PR 24-AUG-2001; 2001US-314821P.
 PR 29-AUG-2001; 2001US-315992P.
 PR 03-MAY-2002; 2002US-378205P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Gandhi AR, Kable AE, Swarnakar A, Hafalia AJA, Tran B, Duggan BM;
 PI Warren BA, Ison CH, Honchell CD, Nguyen DB, Lu DM, Lee EA;
 PI Yue H, Thangaveilu K, Bartoo I, Rankumar J, Griffin JA, Li JX;
 PI Yang J, Forngaveilu K, Gietzen KJ, Ding L, Baughn MR, Borowsky ML;
 PI Yao MG, Wallia NK, Mason PM, Gururajan R, Lee S, Becha SD, Lee SY;
 PI Tran UK, Elliott VS, Luo W, Sprague MW, Tang YT, Lu Y;
 PI Zebajadian Y;
 XX
 DR WPI; 2003-184039/18.
 DR N-PSDB; ABR23206.
 XX
 PT New isolated human PMOD polypeptide and polynucleotide, useful for
 PT diagnosing, treating and preventing diseases or conditions associated
 PT with the aberrant PMOD expression, e.g. cancer, AIDS, atherosclerosis
 PT and infections -
 XX
 PS Claim 62; Page 181-182; 225pp; English.
 XX
 CC The invention relates to an isolated polypeptide comprising: any of 28
 CC sequences of 48-1256 amino acids; a natural amino acid sequence at least
 CC 90% identical to the 28 amino acid sequences, 94% identical to a sequence
 CC of 703 or 267 amino acids, 96% identical to a sequence of 414 amino
 CC acids, or 97% identical to a sequence of 242 amino acids, all given in
 CC the specification; or a biologically active or immunogenic fragment of
 CC the specified polypeptide. The polypeptides and polynucleotides are useful
 CC in diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression of protein modification and maintenance
 CC molecules (PMOD), such as cell proliferative diseases (e.g. cancer,
 CC atherosclerosis), autoimmune/inflammatory diseases (e.g. AIDS,
 CC allergies), neurological disorders (e.g. stroke, Parkinson's disease,
 CC epilepsy), gastrointestinal (e.g. ulcer, cirrhosis), reproductive (e.g.
 CC endometriosis), developmental, vesicle trafficking disorders, and
 CC infections (e.g. bacterial, viral, parasitic, protozoal). These are also
 CC useful in assessing the effects of exogenous compounds on the expression
 CC of nucleic acid and amino acid sequences of PMOD. The PMOD or its
 CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The microarray is useful in monitoring
 CC or measuring protein-protein interactions, drug-target interactions, and
 CC gene expression profiles. This sequence represents a human PMOD protein
 CC of the invention.
 CC
 SO Sequence 545 AA;
 XX
 QY Query Match 87.9%; Score 29; DB 24; Length 545;
 DB Best Local Similarity 83.3%; Pred. No. 7.5e+02;
 DB 411 QOYTD 416 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 OLIT 20
 ABU12117
 ID ABU12117 standard; Protein; 545 AA.
 XX
 AC ABU12117;
 XX

KM systemic infection; non-diarrhoeal infection; septicaemia;
 KM pyelonephritis; antibiotic resistance.
 XX
 OS Escherichia coli.
 XX
 PN WO200166572-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 12-MAR-2001; 2001WO-EP03445.
 XX
 PR 10-MAR-2000; 2000EP-0003145.
 PR 02-FEB-2001; 2001EP-0001449.
 XX
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 PI Bingen E, Bonaccorsi S, Clermont O, Nassif X, Tinsley C;
 DR WPI; 2001-550253/61.
 XX
 PT A library of DNA fragments of Escherichia coli strains for the
 PT phylogenic determination of a given strain comprises polynucleotides of
 PT nature B2/D+ A- .
 XX
 PS Example 6; Fig 6; 646pp; English.
 XX
 CC The invention relates to a library of DNA fragments of Escherichia coli
 CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)
 CC and encoded proteins (ABBS2459-ABBS2919 and ABBS2954-ABBS3094) of nature
 CC B2/D+A-. The polynucleotides have potential anti-inflammatory,
 CC antibacterial and immunosuppressive activity as part of pharmaceutical
 CC compositions used to treat, palliate or prevent extra-intestinal E. coli
 CC infections. The polypeptides are useful for determining the phylogenic
 CC group of a given E. coli strain. These polypeptides can detect and treat
 CC an undesired development of E. coli, particularly an extra-intestinal
 CC infection that include systemic and non-diarrhoeal infections such as
 CC septicaemia, pyelonephritis and meningitis this is particularly
 CC advantageous as bacterial resistance is increasing with the more
 CC frequent use of broad spectrum antibiotic.
 CC
 XX Sequence 1778 AA;
 SQ

Query Match 87.9%; Score 29; DB 22; Length 1778;
 Best Local Similarity 83.3%; Pred. No. 2.6e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQYTDQ 6
 :|||||
 1683 KQYTDQ 1688

RESULT 23
 ABB69513
 ID ABB69513 standard; Protein; 91 AA.
 XX
 AC ABB69513;
 XX
 DT 20-JAN-2003 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1560.
 XX
 KM Human: genome mapping; gene therapy; food supplement; virus; fungus;
 KM cell-proliferative disorder; neurodegenerative disease; bacterial;
 KM Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KM multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KM arthritis; cystostatic; immunomodulator; nootropic; neuroprotective;
 KM antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KM haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 KM antitubercic.
 XX
 OS Homo sapiens.
 XX
 PN WO200270539-A2.

XX
 PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002WO-US05095.
 XX
 PR 05-MAR-2001; 2001US-0799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX
 DR WPI; 2002-759812/82.
 DR N-PSDB; ABZ11730.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
 PT platelet or coagulation disorders -
 XX
 PS Claim 9; SEQ ID NO 1560; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (1) comprising a
 CC nucleotide sequence selected from any of 948 sequences
 CC (ABZ1119-ABZ12066) or their mature protein coding portion, active domain
 CC coding protein or complementary sequences. The polynucleotides are useful
 CC for identifying expressed genes or for physical mapping of human genome.
 CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
 CC weight markers, as a food supplement, for generating antibodies, in
 CC medical imaging, screening and diagnostic assays and for treating
 CC cell-proliferative disorders (cancer), neurodegenerative diseases
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 CC disorders, platelet or coagulation disorders, wound, burns, incision,
 CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
 CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 91 AA;
 SQ

Query Match 84.8%; Score 28; DB 23; Length 91;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQYTDQ 6
 :|||||
 56 QQYTDQ 61

RESULT 24
 AAG02653
 ID AAG02653 standard; Protein; 100 AA.
 XX
 AC AAG02653;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 6734.
 XX
 KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PR 21-FEB-2000; 2000EP-0200610.
 XX

PR 26-FEB-1999; 99US-0122487.
 XX (GEST) GENSET.
 XX
 XX
 PI Dumas Milne-Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-500381/45.
 DR N-PSDB; AAC02659.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 6734; 71bp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 100 AA;
 XX
 Query Match 84.8%; Score 28; DB 21; Length 100;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYTDQ 6
 Db 56 QYTDQ 61
 XX
 DT 08-FEB-2001 (first entry)
 XX
 Human ORFX ORF962 polypeptide sequence SEQ ID NO:1924.
 XX
 KM Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KM vulnery; antiparietic; antiparkinsonian; neurotropic; neuroprotective;
 KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KM antiviral; antibacterial; antifungal; antineumatic; antihydroid;
 KM antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KM thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX

PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkens RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 DR N-PSDB; AAC75407.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 1468-1469; 5507bp; English.
 XX
 CC AAC7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiparietic; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antihydroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 119 AA;
 XX
 Query Match 84.8%; Score 28; DB 21; Length 119;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 QYTDQ 6
 Db 40 QYTDQ 44
 XX
 DT 24-JUN-2002 (first entry)
 XX
 Human ORFX protein sequence SEQ ID NO:17276.
 XX
 KM Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KM hypertension; hypothyroidism; cholesterol ester storage disease;
 KM immune deficiency; immune disorder; infectious disease;
 KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KM myasthenia gravis.
 XX
 OS Homo sapiens.
 XX
 PN WO200192523-A2.
 XX


```

PR 01-JUL-1999; 99US-014842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147036.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

```

```

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 84.8%; Score 28; DB 21; Length 120;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 |||||
 Db 34 QOYTD 38

```

RESULT 28
AAG10235
ID AAG10235 standard; Protein; 130 AA.
XX
AC AAG10235;
XX
DT 17-OCT-2000 (first entry)
XX
DB Arabidopsis thaliana protein fragment SEQ ID NO: 8471.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX

```

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130410.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135533.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 84.8%; Score 28; DB 21; Length 130;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 DB 44 QOYTD 48

RESULT 29

ABBS4752
 ID ABBS4752 standard; Protein; 169 AA.

ABBS4752;

16-MAY-2002 (first entry)

Lactococcus lactis protein p1339.

Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

Lactococcus lactis IL1403.

FR2807446-A1.

12-OCT-2001.

11-APR-2000; 2000FR-0004630.

11-APR-2000; 2000FR-0004630.

(INRG) INRA INST NAT RECH AGRONOMIQUE.

Bolotline A, Sorokine A, Renault P, Ehrlich SD;

WPI; 2002-043418/06.

New nucleotide sequence useful in the identification of Lactococcus

PT lactic and related species -
 XX
 PS Claim 6; SEQ ID NO 1454; 2504bp; French.
 XX
 CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABAS0521) and related proteins (ABBS3300-ABBS5621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 169 AA;

Query Match 84.8%; Score 28; DB 23; Length 169;
 Best Local Similarity 83.3%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTD 6
 DB 64 QOYTD 69

RESULT 30

AAAG10234
 ID AAAG10234 standard; Protein; 175 AA.

AAAG10234;

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 8470.

Protein identification; signal transduction pathway; metabolic pathway;

hybridisation assay; genetic mapping; gene expression control; promoter;

termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.

05-MAR-1999; 99US-0123180.

09-MAR-1999; 99US-0123548.

23-MAR-1999; 99US-0125788.

25-MAR-1999; 99US-0126264.

29-MAR-1999; 99US-0126785.

01-APR-1999; 99US-0127462.

06-APR-1999; 99US-0128234.

08-APR-1999; 99US-0128714.

16-APR-1999; 99US-0129845.

19-APR-1999; 99US-0130077.

21-APR-1999; 99US-0130449.

23-APR-1999; 99US-0130510.

23-APR-1999; 99US-0130891.

28-APR-1999; 99US-0131449.

30-APR-1999; 99US-0132048.

30-APR-1999; 99US-0132407.

04-MAY-1999; 99US-0132484.

05-MAY-1999; 99US-0132485.

06-MAY-1999; 99US-0132486.

07-MAY-1999; 99US-0132487.

11-MAY-1999; 99US-0132863.

14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159854.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 84.8%; Score 28; DB 21; Length 175;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 DB 89 QOYTD 93

LT 31
 8253
 ID AAM28253 standard; Protein; 196 AA.

AC AAM28253;
 DT 07-SEP-1998 (first entry)

DE Amino acid sequence of Uroporphyrin-III C-Methyltransferase.

XX Staphylococcus aureus protein; ribozyme; antisense sequence; control;
 KW Staphylococcal gene; regulatory element; bacterial gene expression;
 KM vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
 KW toxic shock syndrome; Uroporphyrin-III C-Methyltransferase.

XX Staphylococcus aureus.

OS Staphylococcus aureus.

Key Location/Qualifiers
 FH Misc-difference 1..196
 FT "residues designated X are not defined in
 FT the specification"

PN WO9730070-A1.

XX 21-AUG-1997.

XX 19-FEB-1997; 97WO-US02318.

XX 20-FEB-1996; 96US-0011888.

XX (SMIK) SMITHKLINE BEECHAM CORP.

PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO,
 PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;

XX WPI: 1997-424969/39.

XX N-PSDB; AAT84159.

PT Novel polypeptide(s) from Staphylococcus aureus strain WCH29 - used
 PT to isolate antimicrobial compounds, and in vaccines against S.
 PT aureus infection

XX Claim 6; Pages 560-561; 989pp; English.

CC The present sequence represents a Staphylococcus aureus protein, that,
 CC based on homology with a Bacillus megaterium protein, is believed
 CC to be a uroporphyrin-III C-Methyltransferase. The DNA sequence was
 CC isolated from a library of clones of S. aureus WCH 29 in Escherichia
 CC coli. The DNA sequence can be used in the construction of ribozymes and
 CC antisense sequences to control the expression of Staphylococcal genes.
 CC The DNA sequence is also useful as a source of regulatory elements for
 CC the control of bacterial gene expression. The present protein may be
 CC used to produce vaccines to enable a host to produce specific antibodies

CC with antibacterial action. These vaccines and antibodies would protect
 CC a host against invasion by S. aureus, and conditions relating to
 CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled
 CC skin syndrome, and toxic shock syndrome.

SQ Sequence 196 AA;

Query Match 84.8%; Score 28; DB 18; Length 196;
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 DB 57 QOYTD 61

RESULT 32
 ABG60924
 ID ABG60924 standard; Protein; 259 AA.

AC ABG60924;

DT 27-AUG-2002 (first entry)

DE Novel floral meristem identity protein IgMADS.

XX Ryegrass; fescue; MADS-box; MADS-like protein; CENTRORADIALIS;

KW CEN; CEN-like protein; APERIAL2; AP2; AP2-like protein; HB;

KW Homeo-box protein; HB-like protein; plant growth; plant architecture;

KW inflorescence development; flower development; embryo development;

KW seed development; flower organ identity; phase change; male sterility;

KW hybrid seed production; herbage quality; early maturing crop;

KW biomass increase; branching increase; blocking flowering;

KW allergenic pollen; floral meristem identity protein.

XX Lolium perenne.

XX WO200233091-A1.

XX 25-APR-2002.

XX 17-OCT-2001; 2001WO-AU01311.

XX 19-OCT-2000; 2000AU-0000873.

XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.

XX (AGRE-) AGRISearch LTD.

XX Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;

XX WPI: 2002-452388/48.

XX N-PSDB; ABR82075.

XX New substantially purified or isolated polypeptide e.g., MADS-box,

XX CENTRORADIALIS, APERIAL2, Homeo-box proteins, isolated from ryegrass or

XX fescue species, useful for controlling plant life cycles and/or growth

XX phases -

XX Claim 16; Fig 11; 290pp; English.

XX The invention describes a substantially purified or isolated polypeptide

XX (I) from a ryegrass (Lolium sp.) or fescue (Festuca sp.) species, such as

XX MADS-box (MADS) and MADS-like proteins, CENTRORADIALIS (CEN) and CEN-like

CC organ identity through conversion of one floral organ in another.
CC Manipulation of CEN or HB activity in plants alters the control of phase
CC change, flowering time and the number of leaves made before flowering.
CC Manipulation of AP2 activity in plants alters flower organ identity
CC through conversion of one floral organ in another, leads to a change in
CC the number of floral organs and flowering architecture. Manipulation of
CC flowering plant architecture has a wide range of applications such as:
CC inducing male sterility for hybrid seed production; changing flower
CC architecture for enhancing value of ornamentals; delaying flowering in
CC forage grasses thus extending the formation of less digestible stems and
CC increasing herbage quality; altering flowering time allowing early
CC maturing crops; delaying vegetative phase to increase biomass production;
CC increasing branching to enhanced biomass in fruit trees; altering plant
CC size e.g. shorter plant stature; and in blocking flowering and reducing
CC release for allergenic pollen. This is the amino acid sequence of a
CC novel floral meristem identity protein involved in floral development and
CC a potential target for manipulating plant life cycles.

SQ Sequence 259 AA;

Query Match 84.8%; Score 28; DB 23; Length 259;

Best Local Similarity 83.3%; Pred. No. 5.6e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6

Db 209 QSYTDQ 214

RESULT 33

AAU82959

ID AAU82959 standard; Protein; 261 AA.

XX AAU82959;

XX 23-APR-2002 (first entry)

XX Human homologue of BOS1 protein target for antifungal compound.

XX Antifungal; fungal gene transcription; RPP34; POP3; TPA2; NAB2;

XX MPT1; MTR2; BOS1; POL30; RSAC; SQT1; MTW1; TFB1; SPC98; BFR2; RMA1;

XX GCD7; SKI6; NIP1; LCP5; NCE103; ECO1; ORC2; CNS1; YPD1; TIM10; SRB4;

XX yeast; fungus.

XX Homo sapiens.

XX WO200202055-A2.

XX 10-JAN-2002.

XX 28-JUN-2001; 2001WO-US20592.

XX 29-JUN-2000; 2000US-215164P.

XX 10-AUG-2000; 2000US-224457P.

XX (ANAD-) ANADYS PHARM INC.

XX Moore J, Burman ET, Desilva T, Harris S, Komarnitsky S;

XX Mendillo M, Moore D, McCoy M, Sanderson K, Hag T, Zhu S, Long F;

XX Davidov E, Thompson CM;

XX WPI; 2002-147962/19.

XX N-PSDB; ABK32847.

XX Screening candidate antifungal compound for interaction with essential

XX protein, modulation of essential protein activity, binding to essential

XX effects

XX Claim 1; Figure 79; 522pp; English.

XX The invention describes a method of screening a candidate antifungal

XX compound for interaction with essential proteins (EP) or for modulation

CC of EP activity e.g. fungal gene transcription. The proteins tested in the
CC invention include RPP34, POP3, TPA2, NAB2, MPT1, MTR2, BOS1, POL30, RSAC,
CC SQT1, MTW1, TFB1, SPC98, BFR2, RMA1, GCD7, SKI6, NIP1, LCP5, NCE103,
CC ECO1, ORC2, CNS1, YPD1, TIM10 and SRB4 from *S. cerevisiae*; *C. albicans*
CC and human homologues. The method involves contacting a culture with one
CC or more test compounds and determining the effects on the growth or
CC viability of the culture of cells which preferably comprises fungal cells
CC or yeast cells. Preferably the identified compounds interact with, or
CC modulate (preferably inhibit) activity of *C. albicans* EP. The inhibitor
CC compounds identified by the method are useful for preventing or
CC inhibiting fungal, particularly *C. albicans* growth in culture or in a
CC mammal. The antifungal agents interact with essential fungal elements
CC that can be used to treat fungal infection by preventing the growth and
CC preferentially killing the fungi, but does not inhibit the biological
CC activity of mammalian homologues. This amino acid sequence represents a
CC target protein used to test the antifungal compounds, described in the
CC method of the invention.

SQ Sequence 261 AA;

Query Match 84.8%; Score 28; DB 23; Length 261;

Best Local Similarity 83.3%; Pred. No. 5.7e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6

Db 56 QOYTDQ 61

RESULT 34

AAU34323

ID AAU34323 standard; Protein; 312 AA.

XX AAU34323;

XX 14-FEB-2002 (first entry)

XX *Staphylococcus aureus* cellular proliferation protein #599.

XX Antisense; prokaryotic cellular proliferation protein;

XX antibiotic; antibacterial; drug design.

XX *Staphylococcus aureus*.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 22-DEC-2000; 2000US-257931P.

XX 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX N-PSDB; AAS52182.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 5819; 51pp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from Wipo at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 312 AA;
 84.8%; Score 28; DB 22; Length 312;
 at Local Similarity 100.0%; Pred. No. 6.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 QQYTD 5
 187 QQYTD 191
 Db
 RESULT 35
 ABP40664
 ID ABP40664 standard; Protein; 318 AA.
 AC ABP40664;
 XX
 XX 24-JUL-2002 (first entry)
 DT
 XX
 XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5509.
 DE
 XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 KM
 XX Staphylococcus epidermidis.
 OS
 XX US6380370-B1.
 PN
 XX 30-APR-2002.
 PD
 XX 13-AUG-1998; 98US-0134001.
 XX
 XX 14-AUG-1997; 97US-055779P.
 PR 08-NOV-1997; 97US-064964P.
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 XX Doucette-Stamm LA, Bush D;
 PI
 XX MPI: 2002-381255/41.
 DR N-PSDB; ABN93209.
 XX
 XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
 PT polypeptide, useful for diagnosing and treating bacterial infections -
 PR
 XX Disclosure; SEQ ID 5509; 267pp; English.
 PS
 XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP3124 to ABP37960. The *S. epidermidis* sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly *S. epidermidis* infections. The sequences can be used to
 CC screen for compounds able to interfere with the *S. epidermidis* life
 CC cycle or inhibit *S. epidermidis* infection.

CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.
 CC
 SQ Sequence 318 AA;
 84.8%; Score 28; DB 23; Length 318;
 Best Local Similarity 100.0%; Pred. No. 7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 QQYTD 5
 149 QQYTD 153
 Db
 RESULT 36
 AAU33364
 ID AAU33364 standard; Protein; 325 AA.
 AC AAU33364;
 XX
 XX 14-FEB-2002 (first entry)
 DT
 XX Staphylococcus aureus cellular proliferation protein #5.
 DE
 XX Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 KM
 XX Staphylococcus aureus.
 OS
 XX WO200170955-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 21-MAR-2001; 2001WO-US09180.
 PF
 XX 21-MAR-2001; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207272P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 PI
 XX MPI: 2001-611495/70.
 DR N-PSDB; AAS51224.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 PR
 XX Example 3; Seq ID No 3805; 511pp; English.
 PS
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 325 AA;

Query Match 84.8%; Score 28; DB 22; Length 325;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 |||||
 Db 187 QOYTD 191

RESULT 37
 AAU37354 standard; Protein; 325 AA.
 ID AAU37354

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #1524.

KM Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207272P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

PA Hasebebeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ,
 XX Yamamoto RT, Xu HH;

PI WPI; 2001-611495/70.
 N-PSDB; AAS55213.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 12947; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 384 AA;

Query Match 84.8%; Score 28; DB 22; Length 384;
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 |||||
 Db 187 QOYTD 191

RESULT 38
 ABB91551 standard; Protein; 384 AA.
 ID ABB91551

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 762.

KM Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PR (FARB) BAYER AG.

PR Tietjen K, Weidner M;

PR WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -

PS Claim 5; SEQ ID NO 762; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins

XX (ABB90790-ABB94016) for herbicidally active compounds, comprising

XX aligning and comparing nucleic acid or amino acid sequences from plant

XX with nucleic acid or amino acid sequences from non-plant organisms using

XX suitable search parameters, where plant sequences having an E-value

XX greater by a factor of 3 than the E-value of most similar non-plant

XX sequences are selected. The polypeptides or nucleic acids encoding them

XX are useful for identifying modulators. The identified modulators are

XX useful as herbicides.

SQ Sequence 384 AA;

QY 1 QOYTD 6
 |||||
 Db 220 QOYTD 225

RESULT 39
 ABB65406

ID ABB65406 standard; Protein; 447 AA.
 XX
 AC ABB65406;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 23010.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PI 23-MAR-2000; 2000US-191637P.
 XX
 DR 11-JUL-2000; 2000US-0614150.
 XX
 PE (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 XX
 DR N-PSDB; ABL09509.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 23010; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABLU176-ABU30511), expressed DNA
 CC sequences (ABLU1840-ABLU16175) and the encoded proteins
 CC (ABBS7737-ABR72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SO Sequence 447 AA;
 XX
 SQ Query Match 84.8%; Score 28; DB 22; Length 447;
 Best Local Similarity 100.0%; Pred. No. 9.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QQYTD 5
 DB 57 QQYTD 61

RESULT 40
 AAY23206
 ID AAY23206 standard; Protein; 493 AA.
 XX
 AC AAY23206;
 XX
 DT 24-AUG-1999 (first entry)
 XX
 DE Amino acid sequence of Cry3Bb.60 polypeptide.
 XX
 KW Cry3Bb; mutant; insecticidal activity; insecticidal specificity;
 KM coleoptera; southern corn rootworm; western corn root worm;
 KM Diabrotica undecimpunctata howardi Barber; transgenic plant;
 KM Diabrotica virgifera virgifera LeConte; insecticide resistance.
 XX

OS Synthetic.
 OS Bacillus thuringiensis.
 XX
 PN MO9931248-A1.
 XX
 PD 24-JUN-1999.
 XX
 PF 17-DEC-1998; 98WO-US26852.
 XX
 PR 18-DEC-1997; 97US-0996441.
 PR 18-DEC-1997; 97US-0993170.
 PR 18-DEC-1997; 97US-0993722.
 PR 18-DEC-1997; 97US-0993775.
 XX
 PA (ECOG-) ECOGEN INC.
 PA (MONS) MONSANTO CO.
 XX
 PI Bruseck SM, Bryson JW, English L, Kulesza CA, Malvar TM;
 PI Romano C, Slatin SL, Von Terach MA, Walters FS;
 XX
 DR WPI: 1999-395184/33.
 XX
 PT Insecticidal Bacillus thuringiensis proteins
 PT
 PS Claim 39; Page 462-464; 512pp; English.
 XX
 CC AAY23172-Y23206, and AAY23206-X23209 represent new Bacillus
 CC thuringiensis Cry3Bb mutant proteins. The specification also describes
 CC methods of altering Bacillus thuringiensis Cry3Bb. The B. thuringiensis
 CC Cry3Bb polypeptide was modified to have improved insecticidal activity
 CC or enhanced insecticidal specificity against a target insect. The
 CC modification comprises at least one amino acid substitution, addition,
 CC or deletion in the primary sequence of the native or unmodified Cry3Bb
 CC polypeptide, wherein the substitution or deletion occurs at a position
 CC corresponding to from about amino acids 1-365 of the unmodified
 CC polypeptide sequence (AAY23207 represents the wild type Cry3Bb
 CC protein). The polypeptide can be used to kill coleopteran pests,
 CC especially by application to the environment. It is especially
 CC useful against southern corn rootworm and western corn root worm,
 CC (Diabrotica undecimpunctata howardi Barber, and Diabrotica virgifera
 CC virgifera LeConte respectively). The mutant Cry3Bb polynucleotides
 CC can also be used to produce transgenic plants with increased
 CC insecticide resistance.
 CC
 XX
 SO Sequence 493 AA;
 XX
 SQ Query Match 84.8%; Score 28; DB 20; Length 493;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QQYTD 5
 DB 79 QQYTD 83

Search completed: August 20, 2003, 09:29:39
 Job time : 21.4124 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 6.61856 Seconds
(without alignments)
38.357 Million cell updates/sec

Title: US-09-991-433-7

Perfect score: 33

Sequence: 1 QOYTDQ 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

1 number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUTS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	395	3	US-08-856-841-13 Sequence 13, Appl
2	33	100.0	398	3	US-08-856-841-21 Sequence 21, Appl
3	33	100.0	415	3	US-08-856-841-20 Sequence 20, Appl
4	33	100.0	543	3	US-08-856-841-22 Sequence 22, Appl
5	33	100.0	756	4	US-09-438-268-4 Sequence 4, Appl
6	30	90.9	584	4	US-09-328-352-7872 Sequence 7872, Ap
7	29	87.9	640	2	US-08-671-978A-10 Sequence 10, Appl
8	28	84.8	224	4	US-09-107-532A-6761 Sequence 6761, Ap
9	28	84.8	318	4	US-09-134-001C-5509 Sequence 5509, Ap
10	28	84.8	493	3	US-08-996-441B-70 Sequence 70, Appl
11	28	84.8	493	3	US-08-993-170A-70 Sequence 70, Appl
12	28	84.8	493	3	US-08-993-170A-70 Sequence 70, Appl
13	28	84.8	493	3	US-08-993-170A-70 Sequence 70, Appl
14	28	84.8	534	4	US-09-107-532A-6067 Sequence 6067, Ap
15	28	84.8	651	1	US-08-315-468-6 Sequence 6, Appl
16	28	84.8	651	3	US-08-996-441B-52 Sequence 52, Appl
17	28	84.8	651	3	US-08-996-441B-56 Sequence 56, Appl
18	28	84.8	651	3	US-08-996-441B-58 Sequence 58, Appl
19	28	84.8	651	3	US-08-993-722A-52 Sequence 52, Appl
20	28	84.8	651	3	US-08-993-722A-52 Sequence 52, Appl
21	28	84.8	651	3	US-08-993-722A-58 Sequence 58, Appl
22	28	84.8	651	3	US-08-993-170A-52 Sequence 52, Appl
23	28	84.8	651	3	US-08-993-170A-56 Sequence 56, Appl
24	28	84.8	651	3	US-08-993-170A-58 Sequence 58, Appl
25	28	84.8	651	3	US-08-993-775B-52 Sequence 52, Appl
26	28	84.8	651	3	US-08-993-775B-56 Sequence 56, Appl
27	28	84.8	651	3	US-08-993-775B-58 Sequence 58, Appl

28	28	84.8	651	3	US-07-941-650A-4 Sequence 4, Appl
29	28	84.8	652	3	US-08-996-441B-2 Sequence 2, Appl
30	28	84.8	652	3	US-08-996-441B-4 Sequence 4, Appl
31	28	84.8	652	3	US-08-996-441B-6 Sequence 6, Appl
32	28	84.8	652	3	US-08-996-441B-8 Sequence 8, Appl
33	28	84.8	652	3	US-08-996-441B-10 Sequence 10, Appl
34	28	84.8	652	3	US-08-996-441B-12 Sequence 12, Appl
35	28	84.8	652	3	US-08-996-441B-14 Sequence 14, Appl
36	28	84.8	652	3	US-08-996-441B-16 Sequence 16, Appl
37	28	84.8	652	3	US-08-996-441B-18 Sequence 18, Appl
38	28	84.8	652	3	US-08-996-441B-20 Sequence 20, Appl
39	28	84.8	652	3	US-08-996-441B-22 Sequence 22, Appl
40	28	84.8	652	3	US-08-996-441B-24 Sequence 24, Appl
41	28	84.8	652	3	US-08-996-441B-26 Sequence 26, Appl
42	28	84.8	652	3	US-08-996-441B-28 Sequence 28, Appl
43	28	84.8	652	3	US-08-996-441B-30 Sequence 30, Appl
44	28	84.8	652	3	US-08-996-441B-32 Sequence 32, Appl
45	28	84.8	652	3	US-08-996-441B-34 Sequence 34, Appl
					Sequence 36, Appl

ALIGNMENTS

RESULT 1
US-08-856-841-13
Sequence 13, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856, 841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214, 658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US/07/917, 096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 395
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE

HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIOSUM)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
OTHER INFORMATION: mass spectrometry
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
TITLE: GLUTATHIONE S. TRANSFERASE
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
US-08-856-841-13
Query Match 100.0%; Score 33; DB 3; Length 395;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTDQ 6
DB 260 QOYTDQ 265

APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 398
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
OTHER INFORMATION: mass spectrometry
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 21:
US-08-856-841-21

RESULT 2
US-08-856-841-21
Sequence 21, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:

Query Match 100.0%; Score 33; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 249 QOYTDQ 254

RESULT 3

US-08-856-841-20
Sequence 20, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 415
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERTHEMA INFECTION)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURES:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:

PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 20:

US-08-856-841-20

Query Match 100.0%; Score 33; DB 3; Length 415;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 260 QOYTDQ 265

RESULT 4

US-08-856-841-22
Sequence 22, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 543
TYPE: AMINO ACID

TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match 100.0%; Score 33; DB 3; Length 543;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 408 QOYTDQ 413

RESULT 5
US-09-438-268-4
Sequence 4, Application US/09438268
Patent No. 6491907
GENERAL INFORMATION:
APPLICANT: Radinowicz, Joseph E.
APPLICANT: Samulecki, Richard J.
TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
TITLE OF INVENTION: THE SAME
FILE REFERENCE: 5470-186
CURRENT APPLICATION NUMBER: US/09/438,268
CURRENT FILING DATE: 1999-11-10
EARLIER APPLICATION NUMBER: 60/107,840
EARLIER FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 60/123,651
EARLIER FILING DATE: 1999-03-10
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 756
TYPE: PRT
ORGANISM: Virus
US-09-438-268-4

Query Match 100.0%; Score 33; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 408 QOYTDQ 413

DB 601 QOYTDQ 606
RESULT 6
US-09-328-352-7872
Sequence 7872, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7872
LENGTH: 584
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7872

Query Match 90.9%; Score 30; DB 4; Length 584;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 432 QOYTDQ 437

RESULT 7
US-08-671-978A-10
Sequence 10, Application US/08671978A
Patent No. 5959093
GENERAL INFORMATION:
APPLICANT: Saif, Linda J.
APPLICANT: Patwani, Anil
APPLICANT: Kim, Wonyong
APPLICANT: Chang, Keoung-OK
TITLE OF INVENTION: ROTAVIRUS GENES
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CALFEER, HALTER & GRISWOLD
STREET: 800 SUPERIOR AVENUE, SUITE 1400
CITY: CLEVELAND
STATE: OHIO
COUNTRY: USA
ZIP: 44114
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/671,978A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: GOLRICK, MARY E.
REGISTRATION NUMBER: 34,829
REFERENCE/DOCKET NUMBER: 22727/00133
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 622-8200
TELEFAX: (216) 241-0816
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 640 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
US-08-671-978A-10

Query Match 87.9%; Score 29; DB 2; Length 640;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 6
DB 414 QOYTD 419

RESULT 8

US-09-107-532A-6761
Sequence 6761, Application US/09107532A

Patent No. 6583275
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

CORRESPONDENCE ADDRESSES:
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street

CITY: Maltham
STATE: Massachusetts

COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: JULY 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke

REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6761:

SEQUENCE CHARACTERISTICS:
LENGTH: 224 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: YES

ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

FEATURE:
NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...224
SEQUENCE DESCRIPTION: SEQ ID NO: 6761:

Query Match 84.8%; Score 28; DB 4; Length 224;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYTD 6
DB 175 QYTD 179

RESULT 9

US-09-134-001C-5509
Sequence 5509, Application US/09134001C

Patent No. 6380370
GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5509

LENGTH: 318
TYPE: PRT

ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5509

Query Match 84.8%; Score 28; DB 4; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 149 QOYTD 153

RESULT 10

US-08-996-441B-70
Sequence 70, Application US/08996441B

Patent No. 6023013
GENERAL INFORMATION:

APPLICANT: English, Leigh H.
BRUSOCK, Susan M.

APPLICANT: Malvar, Thomas M.
BRYSON, James W.

APPLICANT: Kulesza, Caroline A.
WALTERS, Frederick S.

APPLICANT: Slatin, Stephen L.
VON TERESCH, Michael A.

APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS

NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESSES:

ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433

CITY: Houston
STATE: Texas

COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997

CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:

NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: MECO.151
TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577

INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:

LENGTH: 493 amino acids
TYPE: amino acid

MOLECULE TYPE: linear
MOLECULE TYPE: protein

US-08-996-441B-70

Query Match 84.8%; Score 28; DB 3; Length 493;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 79 QOYTD 83

RESULT 11

US-08-993-722A-70
Sequence 70, Application US/08993722A
Patent No. 6060594

GENERAL INFORMATION:

APPLICANT: English, Leigh H.

APPLICANT: Brussock, Susan M.

APPLICANT: Malvar, Thomas M.

APPLICANT: Bryson, James W.

APPLICANT: Kulesza, Caroline A.

APPLICANT: Walters, Frederick S.

APPLICANT: Slatin, Stephen L.

APPLICANT: Von Tersch, Michael A.

TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED

TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/993,722A

FILING DATE: 18-DEC-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kitchell, Barbara S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: MECO:149

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/474-7577

TELEFAX: 512/474-3106

INFORMATION FOR SEQ ID NO: 70:

SEQUENCE CHARACTERISTICS:

LENGTH: 493 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-993-722A-70

Query Match 84.8%; Score 28; DB 3; Length 493;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5

DB 79 QOYTD 83

RESULT 12

US-08-993-170A-70

Sequence 70, Application US/08993170A

Patent No. 6063597

GENERAL INFORMATION:

APPLICANT: English, Leigh H.

APPLICANT: Brussock, Susan M.

APPLICANT: Malvar, Thomas M.

APPLICANT: Bryson, James W.

APPLICANT: Kulesza, Caroline A.

APPLICANT: Walters, Frederick S.

APPLICANT: Slatin, Stephen L.

APPLICANT: Von Tersch, Michael A.

TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO

TITLE OF INVENTION: COLEOPTERAN INSECTS

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/993,170A

FILING DATE: 18-DEC-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Kitchell, Barbara S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: MECO:002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/474-7577

TELEFAX: 512/474-3106

INFORMATION FOR SEQ ID NO: 70:

SEQUENCE CHARACTERISTICS:

LENGTH: 493 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-993-170A-70

Query Match 84.8%; Score 28; DB 3; Length 493;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5

DB 79 QOYTD 83

RESULT 13

US-08-993-775B-70

Sequence 70, Application US/08993775B

Patent No. 6077824

GENERAL INFORMATION:

APPLICANT: English, Leigh H.

APPLICANT: Brussock, Susan M.

APPLICANT: Malvar, Thomas M.

APPLICANT: Bryson, James W.

APPLICANT: Kulesza, Caroline A.

APPLICANT: Walters, Frederick S.

APPLICANT: Slatin, Stephen L.

APPLICANT: Von Tersch, Michael A.

TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF

TITLE OF INVENTION: DELTA-ENDOTOXINS AGAINST INSECT PESTS

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,775B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:150
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-775B-70

Query Match
Best Local Similarity 84.8%; Score 28; DB 3; Length 493;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 79 QOYTD 83

RESULT 14
US-09-107-532A-6067
Sequence 6067, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-8277
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6067:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1..534
US-09-107-532A-6067

Query Match
Best Local Similarity 84.8%; Score 28; DB 4; Length 534;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTD 6
DB 342 QOYTD 347

RESULT 15
US-08-315-468-6
Sequence 6, Application US/08315468
Patent No. 5554534
GENERAL INFORMATION:
APPLICANT: Michaels, Tracy Ellis
APPLICANT: Poncetrada, Luis
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Process for Controlling Scarab Pests
TITLE OF INVENTION: with Bacillus thuringiensis Isolates
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/315,468
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/014,941
FILING DATE: 01 FEB 1993
APPLICATION NUMBER: 07/828,430
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/808,316
FILING DATE: 16-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA73.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-372-5800
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:

ORGANISM: Bacillus thuringiensis
STRAIN: colworthi
INDIVIDUAL ISOLATE: 43F
IMMEDIATE SOURCE:
CLONE: E. coli XL1-Blue (pML98-4), NRRL B-18291
FEATURE:
NAME/KEY: Protein
LOCATION: 1..651
US-08-315-468-6

Query Match 84.8%; Score 28; DB 1; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 238 QOYTD 242

LT 16
8-996-441B-52
Sequence 52, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-52

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 237 QOYTD 241

RESULT 17
US-08-996-441B-56
Sequence 56, Application US/08996441B
Patent No. 6023013

GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-56

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 237 QOYTD 241

RESULT 18
US-08-996-441B-58
Sequence 58, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS

NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-58

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YTD 5
|||
Db 237 00YTD 241

RESULT 19
US-08-993-722A-52
Sequence 52, Application US/08993722A
Patent No. 6060594
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Teresch, Michael A.
TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,722A
FILING DATE: 18-DEC-1997

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3106
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-722A-52

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YTD 5
|||
Db 237 00YTD 241

RESULT 20
US-08-993-722A-56
Sequence 56, Application US/08993722A
Patent No. 6060594
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Teresch, Michael A.
TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,722A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3106
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-722A-56

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
DB 237 QOYTD 241

RESULT 21

US-08-993-722A-58
; Sequence 58, Application US/08993722A
; Patent No. 6060594
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Tersch, Michael A.
; APPLICANT: Romano, Charles
; TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
; TITLE OF INVENTION: COLPOPTERAN-TOXIC CRYSTAL PROTEINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,722A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3106
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-993-722A-58

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
DB 237 QOYTD 241

RESULT 22

US-08-993-170A-52
; Sequence 52, Application US/08993170A
; Patent No. 6063597
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.

; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Tersch, Michael A.
; TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
; TITLE OF INVENTION: COLPOPTERAN INSECTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,170A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-993-170A-52

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
DB 237 QOYTD 241

RESULT 23

US-08-993-170A-56
; Sequence 56, Application US/08993170A
; Patent No. 6063597
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Tersch, Michael A.
; TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
; TITLE OF INVENTION: COLPOPTERAN INSECTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,170A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
8-993-170A-56

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YTD 5
Db 237 00YTD 241

RESULT 24
US-08-993-170A-58
Sequence 58, Application US/08993170A
Patent No. 6063597
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Terseh, Michael A.
TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
TITLE OF INVENTION: COLEOPTERAN INSECTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESSES:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,170A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:

LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-170A-58

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YTD 5
Db 237 00YTD 241

RESULT 25
US-08-993-775B-52
Sequence 52, Application US/08993775B
Patent No. 6077824
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Terseh, Michael A.
TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF
TITLE OF INVENTION: DELTA-ENDOTOXINS AGAINST INSECT PESTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESSES:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,775B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-775B-52

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YTD 5
Db 237 00YTD 241

RESULT 26
US-08-993-775B-56

Sequence 56, Application US/08993775B
Patent No. 6077824
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,775B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:150
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-775B-56

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 00YTD 5
237 00YTD 241

Db

RESULT 27
US-08-993-775B-58
Sequence 58, Application US/08993775B
Patent No. 6077824
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433

CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,775B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:150
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-775B-58

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 00YTD 5
237 00YTD 241

Db

RESULT 28
US-07-941-650A-4
Sequence 4, Application US/07941650A
Patent No. 6294184
GENERAL INFORMATION:
APPLICANT: Oyeda, Kendrick A.
APPLICANT: Bradfisch, Gregory A.
TITLE OF INVENTION: Process for Controlling Lepidopteran Pests
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/941,650A
FILING DATE: 19920908
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/758,020
FILING DATE: 12-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,935
FILING DATE: 21-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/642,112
FILING DATE: 16-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: tolworthi
INDIVIDUAL ISOLATE: 43f
IMMEDIATE SOURCE:
CLONE: E. coli XL1-Blue (pMT, 98-4), NRRL B-18291
FEATURE:
NAME/KEY: Protein
LOCATION: 1..651
US-07-941-650A-4

Query Match 84.8%; Score 28; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
Db 238 QOYTD 242

RESULT 29
US-08-996-441B-2
Sequence 2, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: von Teresch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-2

Query Match 84.8%; Score 28; DB 3; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
Db 238 QOYTD 242

RESULT 30
US-08-996-441B-4
Sequence 4, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: von Teresch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-4

Query Match 84.8%; Score 28; DB 3; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
Db 238 QOYTD 242

RESULT 31

US-08-996-441B-6
Sequence 6, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas W.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-6

Query Match 84.8%; Score 28; DB 3; Length 652;
at Local Similarity 100.0%; Pred. No. 3.7e+02;
tches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
Db 238 QOYTD 242

RESULT 32
US-08-996-441B-8
Sequence 8, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas W.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-8

Query Match 84.8%; Score 28; DB 3; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
Db 238 QOYTD 242

RESULT 33
US-08-996-441B-12
Sequence 12, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas W.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-12

Query Match 84.8%; Score 28; DB 3; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 238 QOYTD 242

US-08-996-441B-14
Sequence 14, Application US/08996441B
Patent No. 6023013

GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-14

Query Match 84.8%; Score 28; DB 3; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5

DB 238 QOYTD 242

RESULT 35
US-08-996-441B-16
Sequence 16, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-16

Query Match 84.8%; Score 28; DB 3; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 238 QOYTD 242

RESULT 36
US-08-996-441B-18
Sequence 18, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.

APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-18

Query Match 84.8%; Score 28; DB 3; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 238 QOYTD 242

RESULT 37
US-08-996-441B-20
Sequence 20, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B

FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-20

Query Match 84.8%; Score 28; DB 3; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 238 QOYTD 242

RESULT 38
US-08-996-441B-22
Sequence 22, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-22

Query Match 84.8%; Score 28; DB 3; Length 652;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 |||||
 Db 238 QOYTD 242

RESULT 39
 US-08-996-441B-24
 / Sequence 24, Application US/08996441B
 / Patent No. 6023013
 / GENERAL INFORMATION:
 / APPLICANT: English, Leigh H.
 / APPLICANT: Brussock, Susan M.
 / APPLICANT: Malvar, Thomas M.
 / APPLICANT: Bryson, James W.
 / APPLICANT: Kulesza, Caroline A.
 / APPLICANT: Walters, Frederick S.
 / APPLICANT: Slatin, Stephen L.
 / APPLICANT: Von Terssch, Michael A.
 / APPLICANT: Romano, Charles
 / TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
 / NUMBER OF SEQUENCES: 113
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Arnold, White & Durkee
 / STREET: P.O. Box 4433
 / CITY: Houston
 / STATE: Texas
 / COUNTRY: USA
 / ZIP: 77210
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patentin Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/996,441B
 / FILING DATE: 18-DEC-1997
 / CLASSIFICATION: 800
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Kitchell, Barbara S.
 / REGISTRATION NUMBER: 33,928
 / REFERENCE/DOCKET NUMBER: MECO:151
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 512/418-3000
 / TELEFAX: 512/474-7577
 / INFORMATION FOR SEQ ID NO: 24:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 652 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / US-08-996-441B-24

Query Match 84.8%; Score 28; DB 3; Length 652;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 |||||
 Db 238 QOYTD 242

RESULT 40
 US-08-996-441B-26
 / Sequence 26, Application US/08996441B
 / Patent No. 6023013
 / GENERAL INFORMATION:
 / APPLICANT: English, Leigh H.
 / APPLICANT: Brussock, Susan M.
 / APPLICANT: Malvar, Thomas M.

APPLICANT: Bryson, James W.
 APPLICANT: Kulesza, Caroline A.
 APPLICANT: Walters, Frederick S.
 APPLICANT: Slatin, Stephen L.
 APPLICANT: Von Terssch, Michael A.
 APPLICANT: Romano, Charles
 / TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
 / NUMBER OF SEQUENCES: 113
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Arnold, White & Durkee
 / STREET: P.O. Box 4433
 / CITY: Houston
 / STATE: Texas
 / COUNTRY: USA
 / ZIP: 77210
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patentin Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/996,441B
 / FILING DATE: 18-DEC-1997
 / CLASSIFICATION: 800
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Kitchell, Barbara S.
 / REGISTRATION NUMBER: 33,928
 / REFERENCE/DOCKET NUMBER: MECO:151
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 512/418-3000
 / TELEFAX: 512/474-7577
 / INFORMATION FOR SEQ ID NO: 26:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 652 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / US-08-996-441B-26

Query Match 84.8%; Score 28; DB 3; Length 652;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 |||||
 Db 238 QOYTD 242

Search completed: August 20, 2003, 09:34:30
 Job time : 7.61856 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comogen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:30:58 ; Search time 24.3711 Seconds
(without alignments)
32.488 Million cell updates/sec

Title: US-09-991-433-7

Perfect score: 33

Sequence: 1 QOYTDQ 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

1 number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTOS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	756	US-10-205-942-4	Sequence 4, Appli
2	29	87.9	1778	US-10-238-075-749	Sequence 749, App
3	28	84.8	49	US-10-218-102-365	Sequence 365, App
4	28	84.8	112	US-10-079-623-349	Sequence 349, App
5	28	84.8	261	US-09-893-519A-19	Sequence 19, Appli
6	28	84.8	312	US-09-815-242-5819	Sequence 5819, Ap
7	28	84.8	325	US-09-815-242-3805	Sequence 3805, Ap
8	28	84.8	325	US-09-815-242-12947	Sequence 12947, A
9	28	84.8	605	US-10-156-761-14380	Sequence 14380, A
10	28	84.8	652	US-10-233-665-2	Sequence 2, Appli
11	28	84.8	652	US-10-233-665-4	Sequence 4, Appli
12	28	84.8	652	US-10-233-665-6	Sequence 6, Appli
13	28	84.8	653	US-10-233-665-8	Sequence 8, Appli
14	28	84.8	653	US-10-233-665-10	Sequence 10, Appli
15	28	84.8	653	US-10-233-665-12	Sequence 12, Appli

16	28	84.8	653	US-10-232-665-14	Sequence 14, Appli
17	28	84.8	653	US-10-232-665-16	Sequence 16, Appli
18	28	84.8	653	US-10-232-665-18	Sequence 18, Appli
19	28	84.8	653	US-10-232-665-20	Sequence 20, Appli
20	28	84.8	653	US-10-232-665-22	Sequence 22, Appli
21	28	84.8	653	US-10-232-665-24	Sequence 24, Appli
22	28	84.8	653	US-10-232-665-26	Sequence 26, Appli
23	28	84.8	653	US-10-232-665-28	Sequence 28, Appli
24	28	84.8	653	US-10-232-665-30	Sequence 30, Appli
25	28	84.8	653	US-10-232-665-32	Sequence 32, Appli
26	28	84.8	653	US-10-232-665-34	Sequence 34, Appli
27	28	84.8	653	US-10-232-665-36	Sequence 36, Appli
28	28	84.8	653	US-10-232-665-38	Sequence 38, Appli
29	28	84.8	653	US-10-232-665-40	Sequence 40, Appli
30	28	84.8	653	US-10-232-665-42	Sequence 42, Appli
31	28	84.8	653	US-10-232-665-44	Sequence 44, Appli
32	28	84.8	653	US-10-232-665-46	Sequence 46, Appli
33	28	84.8	653	US-10-232-665-48	Sequence 48, Appli
34	28	84.8	653	US-10-232-665-50	Sequence 50, Appli
35	28	84.8	653	US-10-232-665-52	Sequence 52, Appli
36	28	84.8	653	US-10-232-665-54	Sequence 54, Appli
37	28	84.8	653	US-10-232-665-56	Sequence 56, Appli
38	28	84.8	653	US-10-232-665-58	Sequence 58, Appli
39	28	84.8	653	US-10-232-665-60	Sequence 60, Appli
40	28	84.8	653	US-10-232-665-62	Sequence 62, Appli
41	28	84.8	653	US-10-232-665-64	Sequence 64, Appli
42	28	84.8	653	US-10-232-665-66	Sequence 66, Appli
43	28	84.8	653	US-10-232-665-68	Sequence 68, Appli
44	28	84.8	653	US-10-232-665-70	Sequence 70, Appli
45	28	84.8	653	US-10-232-665-72	Sequence 72, Appli

ALIGNMENTS

RESULT 1
US-10-205-942-4
; Sequence 4, Application US/10205942
; Publication No. US20030053990A1
GENERAL INFORMATION:
; APPLICANT: University of No. US20030053990A1th Carolina-Chapel Hill
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/10/205,942
; CURRENT FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Adeno-associated virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(2271)
; OTHER INFORMATION: B19/AAV chimeric capsid coding sequence
US-10-205-942-4
Query Match
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTDQ 6
Db 601 QOYTDQ 606
RESULT 2
US-10-238-075-749
; Sequence 749, Application US/10238075
; Publication No. US20030148324A1

GENERAL INFORMATION:
APPLICANT: I.N.S.E.R.M.
TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolated from E.coli, and biological uses of these polynucleotides and of their derivatives
FILE REFERENCE: BLANDINE
CURRENT APPLICATION NUMBER: US/10/238,075
PRIORITY FILING DATE: 2002-09-10
PRIORITY FILING DATE: 2000-03-10
NUMBER OF SEQ ID NOS: 1576
SOFTWARE: PatentIn version 3.1
SEQ ID NO 749
LENGTH: 1778
TYPE: PRF
ORGANISM: Escherichia coli
US-10-238-075-749

Query Match 87.9%; Score 29; DB 12; Length 1778;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 QOYTDQ 6
:|||||
Db 1683 KOYTDQ 1688

RESULT 3
US-10-218-102-365
Sequence 365, Application US/10218102
Publication No. US20030130827A1
GENERAL INFORMATION:
APPLICANT: Bentzen, Joerg
APPLICANT: Desjardis, John R.
APPLICANT: Hayes, Robert J.
APPLICANT: Velez, Robert J.
TITLE OF INVENTION: Protein Design Automation for Protein Libraries
FILE REFERENCE: A-67229-11/RT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/218,102
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: US 09/927,790
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: US 60/311,545
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: US 60/324,899
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/351,937
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 60/352,103
PRIOR FILING DATE: 2002-01-25
NUMBER OF SEQ ID NOS: 432
SOFTWARE: PatentIn version 3.1
SEQ ID NO 365
LENGTH: 49
TYPE: PRF
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-10-218-102-365

Query Match 84.8%; Score 28; DB 16; Length 49;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 QOYTDQ 6
:|||||
Db 44 QYTDQ 48

RESULT 4
US-10-079-623-349
Sequence 349, Application US/10079623
Publication No. US20020169302A1

GENERAL INFORMATION:
APPLICANT: Havukkala, Ilkka J.
APPLICANT: Glenn, Matthew
APPLICANT: Grigor, Murray R.
APPLICANT: Molenaar, Adrian J.
TITLE OF INVENTION: Compositions isolated from bovine mammary gland and methods for their use.
FILE REFERENCE: 11000,10443
CURRENT APPLICATION NUMBER: US/10/079,623
PRIORITY FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 370
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 349
LENGTH: 112
TYPE: PRF
ORGANISM: Bovine
US-10-079-623-349

Query Match 84.8%; Score 28; DB 14; Length 112;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYTD 5
:|||||
Db 22 QOYTD 26

RESULT 5
US-09-893-519A-19
Sequence 19, Application US/09893519A
Publication No. US20030027243A1
GENERAL INFORMATION:
APPLICANT: ANADYS PHARMACEUTICALS, INC.
APPLICANT: THOMPSON, Craig
APPLICANT: MOORE, Jeffrey
APPLICANT: BURMAN, Ed T.
APPLICANT: BRADLEY, John
APPLICANT: DESIVA, Thamara
APPLICANT: HARRIS, Sandra
APPLICANT: KOMARNITSKY, Svetlana
APPLICANT: MENDILLO, Marc
APPLICANT: MOORE, Daniel
APPLICANT: MCCOY, Melissa
APPLICANT: SANDERSON, Karen
APPLICANT: HAO, Tariq
APPLICANT: ZHU, Shuhao
APPLICANT: LONG, Fan
APPLICANT: DAVIDOV, Eugene
TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
FILE REFERENCE: 0342/IG548-US2
CURRENT APPLICATION NUMBER: US/09/893,519A
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/215,164
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 60/224,457
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 146
SOFTWARE: PatentIn version 3.1
SEQ ID NO 19
LENGTH: 261
TYPE: PRF
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Corresponds to SEQ ID NO: 92
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Human Genbank/JP_003560
DATABASE ENTRY DATE: 2000-11-01
RELEVANT RESIDUES: (1)..(261)
US-09-893-519A-19

Query Match 84.8%; Score 28; DB 11; Length 261;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTD 6
|||:
Db 56 QOYTD 61

RESULT 6
US-09-815-242-5819

; Sequence 5819, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykend, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in
; Prokaryotes

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5819

; LENGTH: 312

; TYPE: PR

; ORGANISM: Staphylococcus aureus

US-09-815-242-5819

Query Match 84.8%; Score 28; DB 9; Length 312;
; Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||:
Db 187 QOYTD 191

RESULT 7
US-09-815-242-3805

; Sequence 3805, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykend, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in
; Prokaryotes

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3805
; LENGTH: 325
; TYPE: PR
; ORGANISM: Staphylococcus aureus
US-09-815-242-3805

Query Match 84.8%; Score 28; DB 9; Length 325;
; Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||:
Db 187 QOYTD 191

RESULT 8
US-09-815-242-12947

; Sequence 12947, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykend, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in
; Prokaryotes

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12947

; LENGTH: 325

; TYPE: PR

; ORGANISM: Staphylococcus aureus

US-09-815-242-12947

Query Match 84.8%; Score 28; DB 9; Length 325;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	QQYTD	5
Db	187	QQYTD	191

RESULT 9
US-10-15

Sequence 14380, Application US/10156761
Publication NO. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARAO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDE
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 14380
LENGTH: 605
TYPE: PAT
ORGANISM: Streptomyces avermitilis
US-10-156-761-14380

Query Match	84.8%	Score 28	DB 15	Length 605
Best Local Similarity	83.3%	Pred. No. 5.2e+02		
Matches	5	Conservative	0	Mismatches 1
				Indels 0
				Gaps 0

QY	1	QQYTDQ	6
Db	72	QQYADQ	77

RESULT 10
US-10-232-665-2

/ Sequence 2, Application US/10232665
 / Publication No. US20030115630A1
 / GENERAL INFORMATION:
 / APPLICANT: Romano, Charles P.
 / TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
 / FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
 / CURRENT APPLICATION NUMBER: US/10/232,665
 / CURRENT FILING DATE: 2002-08-29
 / PRIOR APPLICATION NUMBER: US/09/377,466
 / PRIOR FILING DATE: 1999-08-19
 / NUMBER OF SEQ ID NOS: 43
 / SOFTWARE: PatentIn Ver. 2.0
 / SEQ ID NO 2
 / LENGTH: 652
 / TYPE: PRT
 / ORGANISM: *Bacillus thuringiensis*
 / US-10-232-665-2

Query Match	84.8%;	Score 28;	DB 15;	Length 652;
Best Local Similarity	100.0%;	Pred. NO. 5.6e+02;		
Matches	5;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	1	QQYTD	5
Db	238	QQYTD	242

RESULT 11
US-10-222

US-10-232-665-4
Sequence 4, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Ab Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Ab Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 652
TYPE: PRM
ORGANISM: *Bacillus thuringiensis*
US-10-232-665-4

Query Match	84.8%;	Score 28;	DB 15;	length 652;
Best Local Similarity	100.0%;	Pred. No. 5;	6e+02;	
Matches	5;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0

QY	1	QQYTD	5
Db	238	QQYTD	242

RESULT 12
US-10-232

```

1      Sequence 6, Application US/10232665
2      Publication No. US20030115630A1
3      GENERAL INFORMATION:
4      APPLICANT: Romano, Charles P.
5      TITLE OF INVENTION: Improved Expression of Cry3Ab Insecticidal Protein in Plants
6      FILE REFERENCE: 38-21 (15304) Cry3Ab Improved Exp. Corn
7      CURRENT APPLICATION NUMBER: US/10/232,665
8      CURRENT FILING DATE: 2002-08-29
9      PRIOR APPLICATION NUMBER: US/09/377,466
10     PRIOR FILING DATE: 1999-08-19
11     NUMBER OF SEQ ID NOS: 43
12     SOFTWARE: PatentIn Ver. 2.0
13     SEQ ID NO 6
14     LENGTH: 652
15     TYPE: PRT
16     ORGANISM: Artificial Sequence
17     FEATURE:
18     OTHER INFORMATION: Description of Artificial Sequence: synthetic or
19     FEATURE:
20     OTHER INFORMATION: non-naturally occurring amino acid sequence encoded by SEQ ID NO.1
21     NAME/KEY: PRT
22     LOCATION: (1)..(652)
23     US-10-232-665-6

```

Query Match	84.8%;	Score 28;	DB 15;	Length 652;
Best Local Similarity	100.0%;	Pred. No. 5	6e+02;	
Matches	5;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

QY 1 QQYTD 5
Db 238 QQYTD 242

RESULT 13

US-10-422-865-8
Sequence, 8, Application US/10222665
Publication No. US20030115650A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665

```
/ CURRENT FILING DATE: 2002-08-29
/ PRIOR APPLICATION NUMBER: US/09/377,466
/ PRIOR FILING DATE: 1999-08-19
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 8
/ LENGTH: 653
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: non-naturally
/ OTHER INFORMATION: occurring amino acid sequence encoded by SEQ ID NO:7
/ NAME/KEY: PRT
/ LOCATION: (1)..(653)
/ OTHER INFORMATION: amino acid sequence for Cry3Bb variant v11231 encoded by SEQ ID NO:8
US-10-232-665-8

Query Match
Best Local Similarity 84.8%; Score 28; DB 15; Length 653;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
    |||||
Db 239 QOYTD 243

RESULT 14
US-10-232-665-10
/ Sequence 10, Application US/10232665
/ Publication No. US20030115630A1
/ GENERAL INFORMATION:
/ APPLICANT: Romano, Charles P.
/ TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
/ FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
/ CURRENT APPLICATION NUMBER: US/10/232,665
/ CURRENT FILING DATE: 2002-08-29
/ PRIOR APPLICATION NUMBER: US/09/377,466
/ PRIOR FILING DATE: 1999-08-19
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 10
/ LENGTH: 653
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: non-naturally
/ OTHER INFORMATION: occurring amino acid sequence encoded by SEQ ID NO:9
/ NAME/KEY: PRT
/ LOCATION: (1)..(653)
/ OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO:9
US-10-232-665-10

Query Match
Best Local Similarity 84.8%; Score 28; DB 15; Length 653;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
    |||||
Db 239 QOYTD 243

RESULT 15
US-10-232-665-12
/ Sequence 12, Application US/10232665
/ Publication No. US20030115630A1
/ GENERAL INFORMATION:
/ APPLICANT: Romano, Charles P.
/ TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
/ FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
/ CURRENT APPLICATION NUMBER: US/10/232,665
/ CURRENT FILING DATE: 2002-08-29
```

```
/ PRIOR APPLICATION NUMBER: US/09/377,466
/ PRIOR FILING DATE: 1999-08-19
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 12
/ LENGTH: 653
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: non-naturally
/ OTHER INFORMATION: occurring amino acid sequence encoded by SEQ ID NO:11
/ NAME/KEY: PRT
/ LOCATION: (1)..(1653)
/ OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO:11
US-10-232-665-12

Query Match
Best Local Similarity 84.8%; Score 28; DB 15; Length 653;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
    |||||
Db 239 QOYTD 243

RESULT 16
US-10-232-665-14
/ Sequence 14, Application US/10232665
/ Publication No. US20030115630A1
/ GENERAL INFORMATION:
/ APPLICANT: Romano, Charles P.
/ TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
/ FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
/ CURRENT APPLICATION NUMBER: US/10/232,665
/ CURRENT FILING DATE: 2002-08-29
/ PRIOR APPLICATION NUMBER: US/09/377,466
/ PRIOR FILING DATE: 1999-08-19
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 14
/ LENGTH: 653
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: peptide encoded by SEQ ID NO:
/ NAME/KEY: PRT
/ LOCATION: (1)..(653)
/ OTHER INFORMATION: Cry3Bb1 variant v11231
US-10-232-665-14

Query Match
Best Local Similarity 84.8%; Score 28; DB 15; Length 653;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
    |||||
Db 239 QOYTD 243

RESULT 17
US-10-232-665-16
/ Sequence 16, Application US/10232665
/ Publication No. US20030115630A1
/ GENERAL INFORMATION:
/ APPLICANT: Romano, Charles P.
/ TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
/ FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
/ CURRENT APPLICATION NUMBER: US/10/232,665
/ CURRENT FILING DATE: 2002-08-29
/ PRIOR APPLICATION NUMBER: US/09/377,466
/ PRIOR FILING DATE: 1999-08-19
```

NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 16
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant v11231
US-10-232-665-16

Query Match
Best Local Similarity 100.0%; Score 28; DB 15; Length 653;
Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
DB 239 QOYTD 243

RESULT 18
US-10-232-665-18
Sequence 18, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 18
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant 11231mv1
US-10-232-665-18

Query Match
Best Local Similarity 100.0%; Score 28; DB 15; Length 653;
Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
DB 239 QOYTD 243

RESULT 19
US-10-232-665-20
Sequence 20, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 20
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant 11231mv1
US-10-232-665-20

Query Match
Best Local Similarity 100.0%; Score 28; DB 15; Length 653;
Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
DB 239 QOYTD 243

RESULT 20
US-10-232-665-22
Sequence 22, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 22
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant 11231mv2
US-10-232-665-22

Query Match
Best Local Similarity 100.0%; Score 28; DB 15; Length 653;
Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
DB 239 QOYTD 243

RESULT 21
US-10-232-665-24
Sequence 24, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 24
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant 11231mv2
US-10-232-665-24

Query Match
Best Local Similarity 100.0%; Score 28; DB 15; Length 653;
Pred. No. 5.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
Db 239 QOYTD 243

RESULT 22
US-10-232-665-37
; Sequence 37, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377,466
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PRT
; LOCATION: (1)..(653)
; OTHER INFORMATION: variant Cry3Bb1 coding sequence encoding v11231
US-10-232-665-37

Query Match 84.8%; Score 28; DB 15; Length 653;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
Db 239 QOYTD 243

RESULT 23
US-10-232-665-39
; Sequence 39, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/10/232,665
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PRT
; LOCATION: (1)..(653)
; OTHER INFORMATION: variant Cry3Bb1 coding sequence encoding v11231
US-10-232-665-39

Query Match 84.8%; Score 28; DB 15; Length 653;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
|||||
Db 239 QOYTD 243

RESULT 24
US-09-866-582-35
; Sequence 35, Application US/09866582
; Patent No. US20020127620A1
; GENERAL INFORMATION:
; APPLICANT: Wiltman, George B.
; APPLICANT: Pazour, Gregory J.
; APPLICANT: Rosenbaum, Joel L.
; APPLICANT: Cole, Douglas G.
; TITLE OF INVENTION: INTRACELLULAR TRANSPORT
; FILE REFERENCE: 07917-145001
; CURRENT APPLICATION NUMBER: US/09/866,582
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/206,923
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-866-582-35

Query Match 84.8%; Score 28; DB 10; Length 820;
Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTD 6
|||||
Db 459 QOYTD 464

RESULT 25
US-10-050-704-172
; Sequence 172, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 172
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-050-704-172

Query Match 81.8%; Score 27; DB 15; Length 157;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTD 6
|||||
Db 112 QOYTD 117

RESULT 26
US-09-815-242-5877
; Sequence 5877, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert

```

; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5877
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5877
```

```

Query Match      81.8%; Score 27; DB 9; Length 190;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 QOYTDQ 6
        |||||
Db      6 QOYTDQ 11
```

```

RESULT 27
US-09-815-242-13061
; Sequence 13061, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
```

```

; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13061
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13061
```

```

Query Match      81.8%; Score 27; DB 9; Length 215;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 QOYTDQ 6
        |||||
Db      7 QOYTDQ 12
```

```

RESULT 28
US-09-895-913A-86
; Sequence 86, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; FILE REFERENCE: Encoding No. US20020160456A1 Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-86
```

```

Query Match      81.8%; Score 27; DB 10; Length 225;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 QOYTDQ 6
        |||||
Db     186 QOYTDQ 191
```

```

RESULT 29
US-10-050-704-132
; Sequence 132, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
```

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 132
LENGTH: 267
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (172)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (175)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-050-704-132

Query Match 81.8%; Score 27; DB 15; Length 267;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
|||
222 QOYTDQ 227

RESULT 30
US-10-050-704-297
Sequence 297, Application US/10050704
Publication No. US20030050442A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: 62 Human Secreted Proteins
FILE REFERENCE: P2039P1
CURRENT FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: 09/684,524
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: PCT/US00/08979
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128,693
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 60/130,991
PRIOR FILING DATE: 1999-04-26
NUMBER OF SEQ ID NOS: 344
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 297
LENGTH: 278
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (183)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (186)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-050-704-297

Query Match 81.8%; Score 27; DB 15; Length 278;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
|||
233 QOYTDQ 238

RESULT 31
US-10-156-761-11573
Sequence 11573, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:

APPLICANT: OKURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11573
LENGTH: 404
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-11573

Query Match 81.8%; Score 27; DB 15; Length 404;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
|||
Db 68 QOYTDQ 73

RESULT 32
US-10-197-666A-66
Sequence 66, Application US/1019766A
Publication No. US20030092037A1
GENERAL INFORMATION:

APPLICANT: ASAH KASEI KABUSIKI KAISYA
TITLE OF INVENTION: ELKI phosphorylation related gene
FILE REFERENCE: PH-1548US
CURRENT APPLICATION NUMBER: US/10/197,66A
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: JP 2001-218204
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: JP 2001-263450
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: JP 2002-012176
PRIOR FILING DATE: 2002-01-21
PRIOR APPLICATION NUMBER: US 60/305,884
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/316,304
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 60/350,027
PRIOR FILING DATE: 2002-01-23
NUMBER OF SEQ ID NOS: 156
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 66
LENGTH: 407
TYPE: PRT
ORGANISM: Homo sapiens
US-10-197-666A-66

Query Match 81.8%; Score 27; DB 15; Length 407;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
|||
Db 362 QOYTDQ 367

RESULT 33
US-10-197-666A-142
Sequence 142, Application US/1019766A
Publication No. US20030092037A1
GENERAL INFORMATION:

APPLICANT: ASAH KASEI KABUSIKI KAISYA

;; TITLE OF INVENTION: Elki phosphorylation related gene
;; FILE REFERENCE: PH-1548US
;; CURRENT APPLICATION NUMBER: US/10/197,666A
;; CURRENT FILING DATE: 2002-11-18
;; PRIOR APPLICATION NUMBER: JP 2001-218204
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: JP 2001-263450
;; PRIOR FILING DATE: 2001-08-31
;; PRIOR APPLICATION NUMBER: JP 2002-012176
;; PRIOR FILING DATE: 2002-01-21
;; PRIOR APPLICATION NUMBER: US 60/305,884
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: US 60/316,304
;; PRIOR FILING DATE: 2001-09-04
;; PRIOR APPLICATION NUMBER: US 60/350,027
;; PRIOR FILING DATE: 2002-01-23
;; NUMBER OF SEQ ID NOS: 156
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 142
;; LENGTH: 407
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-197-666A-142

Query Match 81.8%; Score 27; DB 15; Length 407;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 362 QOYTDQ 367

RESULT 34
US-09-815-242-12290
;; Sequence 12290, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Hasebebeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyckind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; TITLE OF INVENTION: Identification of Essential Genes in
;; TITLE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12290
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-815-242-12290

Query Match 81.8%; Score 27; DB 9; Length 469;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 7 QOYTDQ 12

RESULT 35
US-09-824-574-7
;; Sequence 7, Application US/09824574
;; Publication No. US20030077800A1
;; GENERAL INFORMATION:
;; APPLICANT: Rouleau, Natalie
;; APPLICANT: Moilanen, Anu-Maarit
;; APPLICANT: Palvimo, Jorma J.
;; APPLICANT: Jone, Olli A,
;; TITLE OF INVENTION: ARIPI Gene and Protein
;; FILE REFERENCE: 2630-109
;; CURRENT APPLICATION NUMBER: US/09/824,574
;; CURRENT FILING DATE: 2001-04-03
;; NUMBER OF SEQ ID NOS: 7
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 7
;; LENGTH: 2476
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-824-574-7

Query Match 81.8%; Score 27; DB 11; Length 2476;
Best Local Similarity 83.3%; Pred. No. 3.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 2399 QOYTDQ 2404

RESULT 36
US-10-156-761-10200
;; Sequence 10200, Application US/10156761
;; Publication No. US20030119018A1
;; GENERAL INFORMATION:
;; APPLICANT: OKURA, SATOSHI
;; APPLICANT: IKEDA, HARUO
;; APPLICANT: ISHIKAWA, JUN
;; APPLICANT: HORIKAWA, HIROSHI
;; APPLICANT: SHIBA, TADAYOSHI
;; APPLICANT: SAKAKI, YOSHIYUKI
;; APPLICANT: HATTORI, MASAHIRA
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-262
;; CURRENT APPLICATION NUMBER: US/10/156,761
;; CURRENT FILING DATE: 2002-05-29
;; PRIOR APPLICATION NUMBER: JP 2001-204089
;; PRIOR FILING DATE: 2001-05-30
;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 10200
;; LENGTH: 437
;; TYPE: PRT
;; ORGANISM: Streptomyces avermitilis
US-10-156-761-10200

Query Match 78.8%; Score 26; DB 15; Length 437;
Best Local Similarity 83.3%; Pred. No. 9.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 266 QOYTDQ 291

RESULT 37
 US-09-815-915-8
 ; Sequence 8, Application US/09815915
 ; Patent No. US20020025931A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: Meyers, Rachel
 ; TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL
 ; TITLE OF INVENTION: PROTEIN KINASE MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: 38155-20006.00
 ; CURRENT APPLICATION NUMBER: US/09/815,915
 ; CURRENT FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: US 60/191,846
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 1244
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-815-915-8

Query Match 78.8%; Score 26; DB 9; Length 1244;
 Best Local Similarity 66.7%; Pred. No. 2.8e+03;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDO 6
 :|||:
 Db 631 EQYDTR 636

RESULT 38
 US-09-870-759-136
 ; Sequence 136, Application US/09870759
 ; Patent No. US20020177551A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TERMAN, David S
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
 ; FILE REFERENCE: 870759
 ; CURRENT APPLICATION NUMBER: US/09/870,759
 ; CURRENT FILING DATE: 2002-01-14
 ; PRIOR APPLICATION NUMBER: US 60/208,128
 ; PRIOR FILING DATE: 2000-05-30
 ; NUMBER OF SEQ ID NOS: 166
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 136
 ; LENGTH: 1500
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-870-759-136

Query Match 78.8%; Score 26; DB 10; Length 1500;
 Best Local Similarity 66.7%; Pred. No. 3.4e+03;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDO 6
 :|||:
 Db 1137 QOYSDR 1142

RESULT 39
 US-09-826-290-64
 ; Sequence 64, Application US/09826290
 ; Patent No. US2002016468A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Durham, L. Kathryn
 ; APPLICANT: Friedman, David L.
 ; APPLICANT: Herath, Herath Mudiyanselage Achula Chandrasiri
 ; APPLICANT: Kimmel, Lida H.
 ; APPLICANT: Parekh, Rajesh Bhikhu
 ; APPLICANT: Potter, David M.

APPLICANT: Rohlf, Christian
 ; APPLICANT: Silber, B. Michael
 ; APPLICANT: Stiger, Thomas R.
 ; APPLICANT: Sunderland, P. Trey
 ; APPLICANT: Townsend, Robert Reid
 ; APPLICANT: White, Frost
 ; APPLICANT: Williams, Stephen A.
 ; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
 ; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
 ; TITLE OF INVENTION: Alzheimer's Disease
 ; FILE REFERENCE: 2572-1-001 N2
 ; CURRENT APPLICATION NUMBER: US/09/826,290
 ; CURRENT FILING DATE: 2001-04-30
 ; PRIOR APPLICATION NUMBER: US 60/194,504
 ; PRIOR FILING DATE: 2000-04-03
 ; PRIOR APPLICATION NUMBER: US 60/253,647
 ; PRIOR FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 492
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 64
 ; LENGTH: 12
 ; TYPE: PRT
 ; ORGANISM: homo sapien
 US-09-826-290-64

Query Match 75.8%; Score 25; DB 10; Length 12;
 Best Local Similarity 80.0%; Pred. No. 37;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QOYDQ 6
 :|||:
 Db 5 QYDTR 9

RESULT 40
 US-10-058-513-30
 ; Sequence 30, Application US/10058513
 ; Publication No. US20030087245A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Uses of PBH1 in the Diagnosis and Therapeutic Treatment
 ; TITLE OF INVENTION: of Prostate Cancer
 ; FILE REFERENCE: 018501-005910US
 ; CURRENT APPLICATION NUMBER: US/10/058,513
 ; CURRENT FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: US 60/263,951
 ; PRIOR FILING DATE: 2001-01-24
 ; NUMBER OF SEQ ID NOS: 42
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 30
 ; LENGTH: 45
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: transient
 ; OTHER INFORMATION: receptor potential-related channel 7 (TRPC7)
 ; OTHER INFORMATION: peptide sequence
 US-10-058-513-30

Query Match 75.8%; Score 25; DB 15; Length 45;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDO 6
 :|||:
 Db 7 QOYDTR 12

Search completed: August 20, 2003, 10:08:18
 Job time : 25.3711 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 6.74227 Seconds
(without alignments)
85.581 Million cell updates/sec

Title: US-09-991-433-7

Perfect score: 33

Sequence: 1 QOYTDQ 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	781	1 VCPV19	coat protein VP1 -
2	30	90.9	109	2 S77758	probable Ca2+-tran
3	30	90.9	335	2 S52890	hypothetical prote
4	29	87.9	72	2 P00615	ferritin 1 - compe
5	29	87.9	107	1 A32956	chlorodoxin m - Sy
6	29	87.9	271	2 D86767	hypothetical prote
7	29	87.9	310	2 T02202	2'-hydroxyisoflavo
8	29	87.9	415	2 AH0465	anaerobic glycerol
9	29	87.9	595	2 S75439	hypothetical prote
10	29	87.9	727	2 T29612	hypothetical prote
11	29	87.9	736	1 VPXRPC	outer layer protei
12	29	87.9	1588	2 A86036	probable adhesin Z
13	29	87.9	1588	2 H91188	probable adhesin E
14	28	84.8	48	2 T06248	ferritin 1 - soybe
15	28	84.8	130	2 A86415	hypothetical prote
16	28	84.8	169	2 D86801	prophage p13 prote
17	28	84.8	191	2 E64029	hypothetical prote
18	28	84.8	223	2 T19194	hypothetical prote
19	28	84.8	224	2 T19190	hypothetical prote
20	28	84.8	225	2 D81813	hypothetical prote
21	28	84.8	271	2 S64333	hypothetical prote
22	28	84.8	301	2 T35163	probable transcrip
23	28	84.8	309	2 AG0453	homoserine O-succi
24	28	84.8	325	2 G90040	uroporphyrin-III C
25	28	84.8	384	2 G96738	hypothetical prote
26	28	84.8	449	2 D87682	Ompa family protei
27	28	84.8	536	2 T43349	nuclear receptor N
28	28	84.8	652	2 T39811	paraportal crystal
29	28	84.8	659	2 S10228	paraportal crystal

30	28	84.8	851	2 T47305	hypothetical prote
31	28	84.8	1095	2 AC2059	hypothetical prote
32	27	81.8	265	1 S19516	Rv6161 protein - Y
33	27	81.8	426	2 T31187	hypothetical prote
34	27	81.8	469	2 C88960	hypothetical prote
35	27	81.8	484	2 C82171	cardiolipin syntha
36	27	81.8	562	2 C64634	iron-regulated out
37	27	81.8	622	2 B96751	hypothetical prote
38	27	81.8	815	2 F71880	probable iron-regu
39	27	81.8	863	2 JG7517	beta-N-acetylgluco
40	27	81.8	906	2 G75362	aconitate hydratase
41	27	81.8	4558	2 C82199	RTX toxin RtxA Vcl
42	26	78.8	159	2 S52663	ccr protein - Ctr
43	26	78.8	209	2 H85644	hypothetical prote
44	26	78.8	209	2 C90785	probable anti-repr
45	26	78.8	225	2 T16266	hypothetical prote

ALIGNMENTS

RESULT 1
VCPV19
coat protein VP1 - parvovirus B19 (strain Au)
C:Species: parvovirus B19
A>Note: host Homo sapiens (man)
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C:Accession: A24299
R:Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.
J. Virol. 58, 921-936, 1986
A>Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr
A:Reference number: A24299; MUID:86500451; PMID:3701931
A:Accession: A24299
A:Molecule type: DNA
A:Residues: 1-781 <SHA>
A:Cross-References: EMBL:M13178; NID:g333375; PIDN:AAA66867.1; PID:g333377
A:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match
Best Local Similarity 100.0%; Score 33; DB 1; Length 781;
Matches 6; Conservative 0; Pred. No. 19;
Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 626 QOYTDQ 631

RESULT 2
S77758
probable Ca2+-transporting ATPase (EC 3.6.3.8) - Mycoplasma capricolum (fragment)
N:Alternate names: protein MC360
C:Species: Mycoplasma capricolum
C>Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 19-Apr-2002
C:Accession: S77758; S46914
R:Bork, P.; Ouzounis, G.; Casati, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.;
Mol. Microbiol. 16, 955-967, 1995
A>Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physiolo
A:Reference number: S77739; MUID:96059641; PMID:7476192
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-109 <BOR>
A:Cross-References: EMBL:Z33266; NID:9514485; PIDN:CAA83812.1; PID:9530474
A:Experimental source: ATCC 27343
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase
F:1-109/Domain: ATPase nucleotide-binding domain homology (fragment) <ATN>

Query Match
90.9%; Score 30; DB 2; Length 109;

Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|:|:|:|:
Db 7 EQYTDQ 12

RESULT 3

SS2890
hypothetical protein YMR041c - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein YMR532.06c

C/Species: Saccharomyces cerevisiae

C/Date: 08-May-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002

C/Accession: S52890

R/odel, C./ Bowman, S.

Submitted to the EMBL Data Library, February 1995

A/Reference number: S52885

A/Accession: S52890

A/Residues: 1-335 <ODE>

C/Genetics:

A/Gene: MIPS:YMR041c

A/Cross-references: SGD:S0004644

A/Map position: 13R

Query Match
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|:|:|:|:
Db 26 QOYTDQ 31

RESULT 4
PQ0615
ferritin 1 - cowpea (fragment)
C/Species: Vigna unguiculata (cowpea)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-Mar-1995
C/Accession: PQ0615
R/Wicks, R.E.; Entsch, B.
Biochem. Biophys. Res. Commun. 192, 813-819, 1993
A/Title: Functional genes found for three different plant ferritin subunits in the legum
A/Reference number: PQ0613; MUID:93249461; PMID:8484787
A/Accession: PQ0615
A/Molecule type: mRNA
A/Residues: 1-72 <WIC>
A/Cross-references: GB:X67754; GB:X67757
A/Experimental source: leaf, var. M41
C/Superfamily: ferritin
C/Keywords: iron storage

Query Match
Best Local Similarity 87.9%; Score 29; DB 2; Length 72;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|:|:|:|:
Db 25 QOYTDQ 30

RESULT 5

A32956
thioredoxin m - Synechococcus sp.

C/Species: Synechococcus sp.

C/Date: 17-Jul-1992 #sequence_revision 26-May-1994 #text_change 11-Jun-1999

C/Accession: A32956; A30842

R/Muller, E.G.D.; Buchanan, B.B.
J. Biol. Chem. 264, 4008-4014, 1989

A/Title: Thioredoxin is essential for photosynthetic growth. The thioredoxin m gene of A
A/Reference number: A32956; MUID:93139466; PMID:2492995

A/Accession: A32956
A/Molecule type: DNA
A/Residues: 1-107 <MTU>

A/Cross-references: GB:J04475; NID:9142153; PIDN:AAA22057.1; PID:9142154

A/Note: the source is designated as Anacystis nidulans R2, which is also called Synechoc

C/Genetics:

A/Gene: trxm

C/Superfamily: thioredoxin; thioredoxin homology

C/Keywords: redox-active disulfide

F/10-93/Domain: thioredoxin homology <THR>

F/32-35/Disulfide bonds: redox-active #status predicted

Query Match
Best Local Similarity 87.9%; Score 29; DB 1; Length 107;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|:|:|:|:
Db 47 QOYSDQ 52

RESULT 6
D86767
hypothetical protein ylfD [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C/Accession: D86767
R/Bolotin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Welserbach, J.; Ehrli
Genome Res. 11, 731-735, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: D86767
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-271 <STO>

A/Cross-references: GB:AE005176; PID:912724102; PIDN:AAK05238.1; GSPDB:GN00146
A/Experimental source: strain IL1403
C/Genetics:

A/Gene: ylfD

Query Match
Best Local Similarity 87.9%; Score 29; DB 2; Length 271;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|:|:|:|:
Db 237 QOYTDQ 242

RESULT 7
T02202
2'-hydroxyisoflavone reductase (EC 1.3.1.45) - common tobacco
N/Alternate names: protein Ae22
C/Species: Nicotiana tabacum (common tobacco)
C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 21-Jul-2000
C/Accession: T02202
R/Hibi, N.; Higashiguchi, S.; Hashimoto, T.; Yamada, Y.
Plant Cell 6, 723-735, 1994
A/Title: Gene expression in tobacco low-nicotine mutants.

A/Reference number: Z14614; MUID:94312878; PMID:8038607
A/Accession: T02202
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-310 <HIB>

A/Cross-references: EMBL:D28505; NID:9507822; PIDN:BA05866.1; PID:9507823

A/Experimental source: strain Burley 21; tissue-type root

C/Superfamily: 2'-hydroxyisoflavone reductase

C/Keywords: oxidoreductase

Query Match
Best Local Similarity 87.9%; Score 29; DB 2; Length 310;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
|||:
Db 89 QOFTDQ 94

RESULT 8

A:Anaerobic glycerol-3-phosphate dehydrogenase chain C [imported] - Yersinia pestis (strain AH0465)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 03-Nov-2001
C:Accession: AH0465
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AH0465
A>Status: preliminary
A:Molecule type: DNA
A:Cross-references: GB:AL590842; PIND:CA93292.1; PID:g15981739; GSPDB:GN00175
C:Genetics:

A:Gene: glpC
C:Superfamily: glycerol-3-phosphate dehydrogenase (anaerobic) chain C; ferredoxin 2[4Fe-

Query Match 87.9%; Score 29; DB 2; Length 415;
Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
|||:
Db 164 QOYTDQ 169

RESULT 9

A:hypothetical protein slr2105 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S75439
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75439
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-595 <KAN>
A:Cross-references: EMBL:D90911; GB:AB001339; NID:g1653083; PIND:BA18000.1; PID:g165308
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: Synechocystis hypothetical protein slr2105

Query Match 87.9%; Score 29; DB 2; Length 595;
Best Local Similarity 83.3%; Pred. No. 114+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
|||:
Db 290 QOYTDQ 295

RESULT 10

A:hypothetical protein D1014.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29612
R:Du, Z.; Leimbach, D.
submitted to the EMBL Data Library, March 1996

A>Description: The sequence of C. elegans cosmid D1014.
A:Reference number: Z20652
A:Accession: T29612

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-727 <DUN>
A:Cross-references: EMBL:U53180; PIND:AAA96284.1; GSPDB:GN00023; CESP:D1014.5

A:Experimental source: strain Bristol N2; clone D1014

C:Genetics:

A:Gene: CESP.D1014.5
A:Map position: 5
A:Insertions: 52/1; 379/2; 442/2; 579/3; 618/2; 647/3; 676/2; 707/3

Query Match 87.9%; Score 29; DB 2; Length 727;
Best Local Similarity 83.3%; Pred. No. 144+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
|||:
Db 54 QOYSDQ 59

RESULT 11

VXPXPC

outer layer protein VP4 - porcine rotavirus C (strain Cowden)
N:Alternate names: hemagglutinin; outer capsid protein VP4
M:Contains: outer capsid protein VP5; outer capsid protein VP8
C:Species: porcine rotavirus C
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C:Accession: D40822
R:Bremont, M.; Juste-Lesage, P.; Chabanne-Vauterrot, D.; Charpillienne, A.; Cohen, J. Virology 186, 684-692, 1992
A>Title: Sequences of the four larger proteins of a porcine group C rotavirus and compar
A:Reference number: A40822; MUID:92124743; PMID:1310192
A:Accession: D40822
A:Molecule type: genomic RNA
A:Residues: 1-736

A:Cross-references: GB:M74218; NID:g333317; PIND:AB00802.1; PID:g333318

C:Genetics:
A:Map position: segment 4
C:Superfamily: rotavirus outer layer protein VP3
C:Keywords: glycoprotein; hemagglutinin; outer capsid protein
F:1-243/Product: outer capsid protein VP8 #status predicted <VP8>
F:244-250/Region: cleavage processing #status predicted
F:251-736/Product: outer capsid protein VP5 #status predicted <VP5>
F:359,61,64,93,162,191,237,251,304,471,631/Binding site: carbohydrate (Asn) (covalent) #

Query Match 87.9%; Score 29; DB 1; Length 736;
Best Local Similarity 83.3%; Pred. No. 144+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
|||:
Db 417 QOFTDQ 422

RESULT 12

A:probable adhesin Z5029 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
A:Accession: A86036
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: A86036
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glanier, J.D.; Rose, D.J.; Mayhew Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A86036
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1588 <STO>
A:Cross-references: GB:AE005174; NID:g12518349; PIND:AA658749.1; GSPDB:GN00145; UWGP:Z50
A:Experimental source: strain O157:H7, substrain EDL933

C:Gene: 25029

Query Match

Best Local Similarity 87.9%; Score 29; DB 2; Length 1588;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 1493 KOYTDQ 1498

RESULT 13

H91188

probable adhesin ECG4480 [similarity] - *Escherichia coli* (strain O157:H7, substrain R1MD)

C:Species: *Escherichia coli*

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001

C:Accession: H91188

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C. G.

Wara, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

Reg. 8, 11-22, 2001

Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene

Reference number: A99629; MUID:21156231; PMID:11258796

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1588 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA037903.1; PID:G13363955; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:

A:Gene: ECG4480

Query Match
Best Local Similarity 87.9%; Score 29; DB 2; Length 1588;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 1493 KOYTDQ 1498

RESULT 14
T06248

ferritin 1 - soybean (fragment)

C:Species: Glycine max (soybean)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Aug-1999

C:Accession: T06248

R:Kraus, R.E.; Wardrop, A.; Entsch, B.

Submitted to the EMBL Data Library, March 1998

Reference number: Z15565

A:Accession: T06248

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-48 <WIC>

A:Cross-references: EMBL:AF052513; NID:G3003006; PIDN:AC12281.1; PID:G3003008

C:Genetics:

A:Note: Intron positions not resolved (incomplete sequence)

A:Description: Intracellular protein that stores and transports iron in a soluble, nonco

C:Superfamily: ferritin

C:Keywords: iron binding; iron storage; iron transport

Query Match
Best Local Similarity 84.8%; Score 28; DB 2; Length 48;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 13 QOYTDQ 18

RESULT 15
A86415

hypothetical protein F28N24.7 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C:Accession: A86415

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pedersen, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Corn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

anzen, N.F.; Hughes, B.; Hultar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lures, J.S.; Malt, R.; Marshall,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: A86415

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-130 <STO>

A:Cross-references: GB:AB005172; NID:G9502415; PIDN:AAF8114.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match
Best Local Similarity 84.8%; Score 28; DB 2; Length 130;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
DB 44 QOYTD 48

RESULT 16
D86801

prophage p13 protein 39 [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)

C:Species: *Lactococcus lactis* subsp. *lactis*

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: D86801

R:Bolotin, A.; Winkler, P.; Manger, S.; Vallon, O.; Malarne, K.; Weissenbach, J.; Ehrlich

Genome Reg. 11, 731-753, 2001

A>Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: D86801

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-169 <STO>

A:Cross-references: GB:AB005176; PID:G12724400; PIDN:AAK0510.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: p1339

Query Match
Best Local Similarity 84.8%; Score 28; DB 2; Length 169;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 64 QOYTDQ 69

RESULT 17
E64029

hypothetical protein H11422 - *Haemophilus influenzae* (strain Rd KW20)

C:Species: *Haemophilus influenzae*

C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999

C:Accession: E64029

R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.

; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.

; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A>Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: E64029
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-191 <TIGR>
A:Cross-references: GB:U23821; GB:J42023; NID:G1574248; PIDN:AAC23072.1; PID:G1574256; T

Query Match 84.8%; Score 28; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTD 5
DB 35 QOYTD 39

RESULT 18
T19194
hypothetical protein C11E4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19087
A:Accession: T19194
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-223 <MIT>
A:Cross-references: EMBL:Z81015; PIDN:CA802655.1; GSPDB:GN00028; CESP:C11E4.1
A:Experimental source: clone C11E4
C:Genetics:
A:Gene: CESP:C11E4.1
A:Map position: X
A:Introns: 20/1; 44/3; 73/1; 154/3
C:Superfamily: glutathione peroxidase

Query Match 84.8%; Score 28; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTD 5
DB 76 QOYTD 80

RESULT 19
T19180
hypothetical protein C11E4.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19087
A:Accession: T19190
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-224 <MIT>
A:Cross-references: EMBL:Z81015; PIDN:CA802655.1; GSPDB:GN00028; CESP:C11E4.2
A:Experimental source: clone C11E4
C:Genetics:
A:Gene: CESP:C11E4.2
A:Map position: X
A:Introns: 22/1; 74/1; 155/3
C:Superfamily: glutathione peroxidase

Query Match 84.8%; Score 28; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTD 5
DB 77 QOYTD 81

RESULT 20

D81813
hypothetical protein NMA1865 (imported) - Neisseria meningitidis (strain Z2491 serogroup
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: D81813
R:Parthill, J.; Achtmann, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: D81813
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-225 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:G7380371; PIDN:CA85088.1; PID:G738050
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1865

Query Match 84.8%; Score 28; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTD 5
DB 169 QOYTD 173

RESULT 21

SC4333
hypothetical protein YGR042w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G4154
C:Species: Saccharomyces cerevisiae
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
A:Accession: SC4333; SC4336
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: SC4333
A:Molecule type: DNA
A:Residues: 1-271 <RIB>
A:Cross-references: EMBL:Z72827; NID:G1323040; PIDN:CAA97040.1; PID:E243403; PID:G132304

A:Experimental source: strain S288C
A:Status: preliminary
A:Reference number: S64335
A:Accession: SC4336
A:Molecule type: DNA
A:Residues: 1-271 <BNT>
A:Cross-references: EMBL:Z72827; NID:G1323040; PIDN:CAA97040.1; PID:E243403; PID:G132304

A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0003274
A:Map position: 7R

Query Match 84.8%; Score 28; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 QYTD 6
DB 11 QYTD 15

RESULT 22

T35163
probable transcription regulator - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35163

R. Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1998
 A:Reference number: Z21570
 A:Accession: T335163
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-301 <SEE>
 A:Cross-references: EMBL:AL031107; PIDN:CAA19943.1; GSPDB:GN00070; SCOEDB:SC5A7.14
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC5A7.14

Query Match 84.8%; Score 28; DB 2; Length 301;
 Best Local Similarity 83.3%; Pred. No. 91;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 286 OQYTDQ 291

RESULT 23

AG0453
 homoserine O-succinyltransferase (EC 2.3.1.46) [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AG0453
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchell, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11, M.; Rutherford, K.; Skellern, M.; Skellern, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AG0453
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-309 <RUR>
 A:Cross-references: GB:AL590842; PIDN:CAC93195.1; PID:gl5981643; GSPDB:GN00175
 C:Genetics:
 A:Gene: meta
 C:Superfamily: homoserine succinyltransferase
 C:Keyword: acyltransferase; coenzyme A

Query Match 84.8%; Score 28; DB 2; Length 309;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYTD 5
 203 OQYTD 207

RESULT 24
 G90040
 utroporphyrin-III C-methyl transferase [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: G90040
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: AB9758; MUID:21311952; PMID:11418146
 A:Accession: G90040
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-325 <RUR>
 A:Cross-references: GB:BA000018; PID:gl3702347; PIDN:BA843488.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: nasf

Query Match 84.8%; Score 28; DB 2; Length 325;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 DB 187 OQYTD 191

RESULT 25

G96738
 hypothetical protein F14023.4 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: G96738
 R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Corn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marshall,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A66141; MUID:21016719; PMID:11130712
 A:Accession: G96738
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-384 <STO>
 A:Cross-references: GB:AE005173; NID:G7239493; PIDN:AAF43219.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F14023.4
 C:Superfamily: myrosinase-associated protein MyAP

Query Match 84.8%; Score 28; DB 2; Length 384;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 220 OQYTDQ 225

RESULT 26

DB7682
 Ompa family protein [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: DB7682
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.U.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 u, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete genome sequence of Caulobacter crescentus.
 A:Reference number: AB7249; MUID:21173698; PMID:11259647
 A:Accession: DB7682
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-449 <STO>
 A:Cross-references: GB:AE005673; NID:gl3425220; PIDN:AAK25456.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC3494

Query Match 84.8%; Score 28; DB 2; Length 449;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYTDQ 6
 DB 289 OYTDQ 293

RESULT 27

T43349

nuclear receptor NHR-7 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C:Accession: T43349; T22642

R:Slinder, A.E.; Mathews, S.W.; Hough, D.; Yin, V.P.; Maina, C.V.

Genome Res. 9, 103-120, 1999

A:Title: The nuclear receptor superfamily has undergone extensive proliferation and divergence

A:Reference number: Z22443; MUID:99148134; PMID:10022975

A:Accession: T43349

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-536 <SLU>

A:Cross-references: EMBL:AF083225; NID:g4139075; PIDN:AAD03683.1; PID:g4139076

R:Lenhard, N.

Submitted to the EMBL Data Library, July 1996

A:Reference number: Z19592

A:Accession: T22642

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-536 <ML>

A:Cross-references: EMBL:T77132; PIDN:CAB00862.1; GSPDB:GN00022; CESP:F54D1.4

A:Experimental source: clone F54D1

C:Genetics:

A:Gene: nhr-7; F54D1.4

A:Map position: 4

A:introns: 18/3; 57/1; 79/1; 215/3; 255/2; 293/1; 346/3

Query Match

Best Local Similarity 84.8%; Score 28; DB 2; Length 536;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5

Db 439 QOYTD 443

RESULT 28

T39811

paraoporal crystal protein cry3Bb1 - *Bacillus thuringiensis*

N:Alternate names: paraoporal crystal protein cryIIIB2

C:Species: *Bacillus thuringiensis*

C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 01-Dec-2000

C:Accession: T39811

R:Donovan, W.P.; Rupp, M.J.; Slaney, A.C.; Malvar, T.; Gawron-Burks, M.C.; Johnson, T.H.

J. Biol. Environ. Microbiol. 58, 3921-3927, 1992

A:Title: Characterization of two genes encoding *Bacillus thuringiensis* insecticidal cry

A:Reference number: T39811; MUID:93119147; PMID:1476436

A:Accession: T39811

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-652 <RES>

A:Cross-references: GB:M89794; NID:G142729; PIDN:AAA2234.1; PID:G142730

C:Genetics:

A:Gene: cryIIIB2

C:Superfamily: paraoporal crystal protein

Query Match

Best Local Similarity 84.8%; Score 28; DB 2; Length 652;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5

Db 238 QOYTD 242

RESULT 29

S10228

paraoporal crystal protein cry3Bb1 - *Bacillus thuringiensis* (fragment)

N:Alternate names: coleopteran-active paraoporal crystal protein; delta-endotoxin

C:Species: *Bacillus thuringiensis*

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Dec-2000

C:Accession: S10228

R:Sick, A.; Gaertner, F.; Wong, A.

Nucleic Acids Res. 18, 1305, 1990

A:Title: Nucleotide sequence of a coleopteran-active toxin gene from a new isolate of *B*

A:Reference number: S10228; MUID:90206811; PMID:2320431

A:Accession: S10228

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-659 <SIC>

A:Cross-references: EMBL:X17123; NID:g40258; PIDN:CAA34983.1; PID:g40259

C:Genetics:

A:Gene: cryIIIB

C:Superfamily: paraoporal crystal protein

Query Match

Best Local Similarity 84.8%; Score 28; DB 2; Length 659;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5

Db 246 QOYTD 250

RESULT 30

T47305

hypothetical protein T32A11.40 - *Arabidopsis thaliana*C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47305

R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, I.

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24459

A:Accession: T47305

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-851 <RIE>

A:Cross-references: EMBL:AL138653

A:Experimental source: cultivar Columbia; BAC clone T32A11

C:Genetics:

A:Map position: 3

A:introns: 47/3; 69/3; 293/3; 355/2; 372/1; 418/1; 557/3; 584/3; 620/3; 644/3; 681/3; 71

A>Note: T32A11.40

Query Match

Best Local Similarity 84.8%; Score 28; DB 2; Length 851;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 6

Db 487 QOYTD 492

RESULT 31

AC2059

hypothetical protein al12025 [imported] - *Nostoc* sp. (strain PCC 7120)C:Species: *Nostoc* sp. PCC 7120A>Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AC2059

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Matsumoto, A.; Iritani, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium *Ana*

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AC2059

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1095 <KTR>

A:Cross-references: GB:BA000019; PIDN:BA073724.1; PID:g17131116; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: a112025

Query Match 84.8%; Score 28; DB 2; Length 1095;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 5
 |||||
 DB 206 QOYTD 210

RESULT 32

S19516

RV161 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YCR009c

C:Species: Saccharomyces cerevisiae

C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 21-Jul-2000

C:Accession: S19516; S28651; S22849

R:Goffeau, A.; Purnelle, B.; Skala, J.

Submitted to the Protein Sequence Database, March 1992

Reference number: S19420

C:Accession: S19516

A:Molecule type: DNA

A:Residues: 1-265 <GOF>

A:Cross-references: EMBL:X59720; NID:G1907116; PIDD:CAA42326.1; PID:G1907152; GSPDB:GN00

R:Crouzet, M.; Urdaci, M.; Dulau, L.; Aigle, M.

Yeast 7, 727-743, 1991

A:Title: Yeast mutant affected for viability upon nutrient starvation: characterization

A:Reference number: S28651; MUID:9213163; PMID:1776363

C:Accession: S28651

A:Molecule type: DNA

A:Residues: 1-94, 'E', 96-265 <CRO>

A:Cross-references: EMBL:X63315; NID:G4417; PIDD:CAA44926.1; PID:G4418

R:Urdaci, M.; Dulau, L.; Aigle, M.; Crouzet, M.

Yeast 6, 173-176, 1990

A:Title: Sequence of the yeast gene RV5161 located on chromosome III.

A:Reference number: S22849; MUID:90224366; PMID:2183524

C:Accession: S22849

A:Molecule type: DNA

A:Residues: 1-94, 'E', 96-265 <URD>

A:Cross-references: EMBL:X63315; NID:G4417; PIDD:CAA44926.1; PID:G4418

R:Skala, J.; Purnelle, B.; Goffeau, A.

Yeast 8, 409-417, 1992

A:Title: The complete sequence of a 10.8 kb segment distal of SUR2 on the right arm of C

X genes.

A:Reference number: S25353; MUID:92327849; PMID:1626432

A:Contents: annotation

A:Genetics:

A:Map position: 3R

A:Superfamily: RV5161 protein; RV5161 protein homology

F:4-257/Dmain: RV5161 protein homology <RVS>

Query Match 81.8%; Score 27; DB 1; Length 265;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 |||||
 DB 227 QOYTDQ 232

RESULT 33

T33187

hypothetical protein F22F7.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Aug-2000

C:Accession: T33187

R:Miller, N.; Kramer, J.; Smith, A.

Submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid F22F7.

A:Reference number: Z21299

A:Accession: T33187
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-426 <ML>

A:Cross-references: EMBL:AF067937; PIDD:AAC19221.1; GSPDB:GN00023; CESP:F22F7.1

A:Experimental source: strain Bristol N2; clone F22F7

C:Genetics:

A:Gene: CESP:F22F7.1

A:Map position: 5

A:Introns: 56/1; 91/3; 232/1; 392/2

A:Superfamily: Streptomyces coelicolor conserved hypothetical protein SC5F2A.12c

Query Match 81.8%; Score 27; DB 2; Length 426;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 |||||
 DB 304 QOYTDQ 309

RESULT 34

C89960

hypothetical protein SA1572 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: C89960

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc

ma, A.; Mizutani-Ut, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratake, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

C:Accession: C89960

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-469 <KUR>

A:Cross-references: GB:BA000018; PID:G13701547; PIDD:BA842840.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA1572

C:Superfamily: peptidase V

Query Match 81.8%; Score 27; DB 2; Length 469;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 |||||
 DB 7 QOYTDQ 12

RESULT 35

C82171

cardiolipin synthase VC1670 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: C82171

R:Heideberg, U.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

Chardoon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers, P

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

C:Accession: C82171

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-484 <HEI>

A:Cross-references: GB:AE004245; GB:AE003852; NID:G9656183; PIDD:AAF94820.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961, biotype El Tor

C:Genetics:

A:Gene: VC1670

A:Map position: 1

C:Superfamily: Bacillus probable cardiolipin synthetase

Query Match 81.8%; Score 27; DB 2; Length 484;

Best Local Similarity 83.3%; Pred. No. 2.6e+02; Mismatches 1; Indels 0; Gaps 0;

DB 1 QOYTDQ 6
449 QOYTDQ 454

RESULT 36

C64634

Iron-regulated outer membrane protein - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: C64634

R:From: J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodok, A.; McKenna

J.D.; Kelley, J.M.; Cotton, M.D.; Weldman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.

Re 388, 539-547, 1997

R:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: C64634

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-562 <TOM>

A:Cross-references: GB:AE00601; GB:AE000511; NID:g2314051; PIDN:AAD07961.1; PID:g231405

Query Match 81.8%; Score 27; DB 2; Length 562;

Best Local Similarity 83.3%; Pred. No. 3.1e+02; Mismatches 1; Indels 0; Gaps 0;

DB 1 QOYTDQ 6
523 QOYTDQ 528

RESULT 37

Hypothetical protein F28P22.16 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: B96751

R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Comn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

en, N.F.; Hughes, B.; Huizart, L.

Re 408, 816-820, 2000

R:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lian, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A66141; MUID:21016719; PMID:11130712

A:Accession: B96751

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-622 <STO>

A:Cross-references: GB:AE005173; NID:g6648163; PIDN:AAF21163.1; GSPDB:GN00141

C:Genetics:

A:Gene: F28P22.16

A:Map position: 1

Query Match 81.8%; Score 27; DB 2; Length 622;

Best Local Similarity 83.3%; Pred. No. 3.4e+02; Mismatches 1; Indels 0; Gaps 0;

DB 1 QOYTDQ 6
134 QOYTDQ 139

RESULT 38

F71880

probable iron-regulated outer membrane protein - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Sep-1999

C:Accession: F71880

R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: F71880

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-815 <ARN>

A:Cross-references: GB:AE001514; GB:AE001439; NID:g4155418; PIDN:AAD06429.1; PID:g415542

A:Experimental source: strain J99

C:Genetics:

A:Gene: frpB_2

C:Superfamily: iron-regulated outer membrane protein

Query Match 81.8%; Score 27; DB 2; Length 815;

Best Local Similarity 83.3%; Pred. No. 4.6e+02; Mismatches 1; Indels 0; Gaps 0;

DB 1 QOYTDQ 6
776 QOYTDQ 781

RESULT 39

JC7537

beta-N-acetylglucosaminidase - Alteromonas sp. (strain O-7)

C:Species: Alteromonas sp. (strain O-7)

C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 01-Mar-2002

C:Accession: JC7537; PC7108

R:Tsujiho, H.; Miyamoto, J.; Kondo, N.; Miyamoto, K.; Baba, N.; Inamori, Y.

Biosci. Biotechnol. Biochem. 64, 2512-2516, 2000

A:Title: Molecular cloning of the gene encoding an outer-membrane-associated beta-N-acet

A:Reference number: JC7537; MUID: 21036923; PMID:11193430

A:Accession: JC7537

A:Molecule type: DNA

A:Residues: 1-863 <TSU>

A:Cross-references: DDBJ:AB042622

A:Experimental source: strain O-7

A:Accession: PC7108

A:Molecule type: protein

C:Comment: This enzyme, an outer-membrane-associated lipoprotein, which belongs to the f

glucosamine.

C:Genetics:

A:Gene: GLCNACaseA

C:Superfamily: beta-hexosaminidase

C:Keywords: lipoprotein

Query Match 81.8%; Score 27; DB 2; Length 863;

Best Local Similarity 83.3%; Pred. No. 4.9e+02; Mismatches 1; Indels 0; Gaps 0;

DB 1 QOYTDQ 6
831 QOYTDQ 836

RESULT 40

G75362

aconitate hydratase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: G75362

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: G75362
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-906 <WHI>
A/Cross-references: GB:AE002013; GB:AE000513; NID:g6459484; PIDV:AAF11276.1; PID:g645949
A/Experimental source: strain R1
C/Genetics:
A/Gene: DRI720
A/Map position: 1
C/Superfamily: Iron-responsive element-binding protein

Query Match 81.8%; Score 27; DB 2; Length 906;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 QOYTDQ 6
|||
DB 492 QOYTDQ 497

Search completed: August 20, 2003, 09:32:42
Job time : 8.74227 secs

DE Hypothetical 38.2 kDa protein in SUB1-ARG1 intergenic region.
GN YMR041C OR YMR532.06C.
OS Saccharomyces cerevisiae (Baker's Yeast).
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=5286C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagsels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RL Nature 387:90-93(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 248502; CAA88407.1; -.
CC PIR; S52890; S52890.
CC DR SGD; S0004644; YMR041C.
CC DR InterPro; IPR001395; Aldo/ket_red.
CC DR Pfam; PF00248; Aldo_ket_red; 1.
CC DR ProDom; PD000288; Aldo/ket_red; 1.
CC KW Hypothetical protein.
SQ SEQUENCE 335 AA; 38220 MW; CC7FA464FPA6ED7 CRC64;
Query Match 90.9%; Score 30; DB 1; Length 335;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTDQ 6
Db 26 QOYTDQ 31
RESULT 3
TX6 BRARE STANDARD; PRT; 473 AA.
P79742;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX6 (T-box protein 6).
GN TBX6.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
CC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=97236936; PubMed=9119115;
RA Hug B., Walter V., Grunwald D.J.;
RT "tbx6, a Brachyury-related gene expressed by ventral mesodermal
precursors in the zebrafish embryo.";
RL Dev. Biol. 183:61-73(1997).
CC -----
CC -1- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
DEVELOPMENTAL PROCESSES.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: EXCLUSIVELY EXPRESSED BY VENTRAL MESENODERM.
CC -1- SIMILARITY: Contains 1 T-box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U80951; AAB39319.1; -.
CC DR HSSP; P24781; 1XR.
CC DR ZFIN; ZDB-GENE-980526-171; tbx6.
CC DR InterPro; IPR001699; TF_T-box.
CC DR Pfam; PF00907; T-box; 1.
CC DR PRINTS; PR00937; TBOX.
CC DR SMART; SM00425; TBOX; 1.
CC DR PROSITE; PS01283; TBOX_1; 1.
CC DR PROSITE; PS01264; TBOX_2; 1.
CC DR PROSITE; PS0252; TBOX_3; 1.
CC KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA BIND 43 217
T-BOX.
SQ SEQUENCE 473 AA; 53477 MW; 10A5066D02AFAD2B CRC64;
Query Match 90.9%; Score 30; DB 1; Length 473;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTDQ 6
Db 448 QOYTDQ 453
RESULT 4
TX1 SYN7 STANDARD; PRT; 106 AA.
ID TX1 SYN7
AC P12243;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thioredoxin 1 (TRX-1) (Thioredoxin M).
GN TRXA OR TRXM.
OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
CC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1140;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89139466; PubMed=249295;
RA Muller E.G.D., Buchanan B.B.;
RT "Thioredoxin is essential for photosynthetic growth. The thioredoxin
m gene of Anacystis nidulans.";
RL J. Biol. Chem. 264:4008-4014(1989).
CC -1- FUNCTION: Participates in various redox reactions through the
reversible oxidation of its active center dithiol to a disulfide
and catalyzes dithiol-disulfide exchange reactions.
CC -1- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04475; AAA22057.1; -.
CC DR HSSP; P80579; 1OYW.
CC DR InterPro; IPR006662; ThioRed.
CC DR InterPro; IPR006663; ThioRedox_dom2.
CC DR InterPro; IPR005746; Thioredoxin.
CC DR Pfam; PF00085; thioRed; 1.
CC DR PRINTS; PR00421; THIOREDOXIN.
CC DR TIGRPFAMs; TIGR01068; thioredoxin; 1.
CC DR PROSITE; PS00194; THIOREDOXIN; 1.
CC KW Redox-active center; Electron transport.
FT INT MET 0
FT DISULFID 31 34
REDOX-ACTIVE (BY SIMILARITY).

```

RA MEDLINE=92124743; PubMed=1310192;
RA Brumont M., Juste-Lesage P., Chabanne-Vautherot D.,
RA Charpilletienne A., Cohen U.;
RT "Sequences of the four larger proteins of a porcine group C rotavirus
RT and comparison with the equivalent group A rotavirus proteins";
RL Virology 186:1684-692(1992).
CC -1- SUBCELLULAR LOCATION: Outer capsid.
CC CC
CC -1- PM: VP8 IS ONE OF TWO TRIPSPIN CLEAVAGE PRODUCTS OF VP4; THE
CC OTHER PRODUCT IS VP5.
CC -1- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL, M74218; AAB00802.1; -.
DR PIR, D40822; VPKRPC.
DR InterPro; IPR000416; Cap_VP4.
DR Pfam; PF00426; VP4; 1.
KW Coat protein; Glycoprotein..
FT CHAIN 1 736 OUTER CAPSID PROTEIN VP4.
FT CHAIN 1 244 OUTER CAPSID PROTEIN VP8 (POTENTIAL).
FT CARBOHYD 39 39 OUTER CAPSID PROTEIN VP5 (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 736 AA; 83231 MW; 356F9226D5016577 CRC64;
Query Match 87.9%; Score 29; DB 1; Length 736;
Best Local Similarity 83.3%; Pred. No. 75;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QQYTDQ 6
Db 417 QQFTDQ 422

RESULT 7
ID Z02 MOUSE STANDARD; PRT; 1167 AA.
AC Q9Z0U1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tight junction protein ZO-2 (Zonula occludens 2 protein) (Zona
DE occludens 2 protein) (Tight junction protein 2).
GN Tjp2 OR ZO2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA;
RX MEDLINE=99150392; PubMed=10026224;
RX Itoh M., Morita K., Teukita S.;
RT "Characterization of ZO-2 as a MAGUK family member associated with
RT tight as well as adherens junctions and a binding affinity to
RT occludin and alpha catenin.";
RL J. Biol. Chem. 274:5981-5986(1999).
CC -1- FUNCTION: PLAYS A ROLE IN TIGHT JUNCTIONS AND ADHERENS JUNCTIONS.
CC

```

```

CC -1- SUBUNIT: INTERACTS WITH OCCULUDIN.
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE, CYTOPLASMIC SIDE.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -1- SIMILARITY: Contains 3 PDZ/DRH domains.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -1- SIMILARITY: Contains 1 guanylate kinase-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF113005; AAD19964.1; -.
DR HSSP; P31016; 1B89.
DR MGD; MGI:1341872; TjP2.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR InterPro; IPR000619; Guanylate_kin.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR005417; Zonoccludens.
DR InterPro; IPR005419; Zonoccluds2.
DR Pfam; PF00625; Guanylate_kin; 1.
DR Pfam; PF00595; PDZ; 3.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01597; ZONOCCLUDNS.
DR PRINTS; PR01599; ZONOCCLUDNS2.
DR SMART; SM00072; GUKC; 1.
DR SMART; SM00328; PDZ; 3.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; FALSE_NEG.
DR PROSITE; PS00852; GUANYLATE_KINASE_2; 1...
DR PROSITE; PS50106; PDZ; 3.
DR PROSITE; PS50002; SH3; 1.
DR TIGR; H11422; SH3 domain; Repeat; Membrane.
DR FT DOMAIN 10 97 PDZ 1.
DR FT DOMAIN 287 365 PDZ 2.
DR FT DOMAIN 489 570 PDZ 3.
DR FT DOMAIN 584 649 SH3.
DR FT DOMAIN 678 858 GUANYLATE_KINASE.
DR FT DOMAIN 1139 1142 POLY-GLU.
DR SQ SEQUENCE 1167 AA; 131614 MW; F15DA3EBC3P9434F CRC64;

Query Match 87.9%; Score 29; DB 1; Length 1167;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 QOYTDQ 6
401 QOYSDQ 406

RESULT 8
YE22 HAEIN STANDARD; PRT; 191 AA.
ID YE22 HAEIN
AC P44153;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein H11422.
GN H11422.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellales; Haemophilus.
OC NCB1_TaxID=727;
OX [1]
RN [1]
RP SEQUENCE FROM N.A. ATCC 51907;
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PUBMED=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

```

```

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodex A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Pine L.D., Fleischmann J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32821; AAC33072.1; -.
DR PIR; E64029; E64029.
DR TIGR; H11422; -.
DR KW Hypothetical protein; Complete proteome.
DR SQ SEQUENCE 191 AA; 22701 MW; D1F632B6066E785 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 191;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
Db 35 QOYTD 39

RESULT 9
GSHU CAEEL STANDARD; PRT; 224 AA.
ID GSHU CAEEL
AC Q95003;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Glutathione peroxidase precursor (EC 1.11.1.9).
GN C1B4.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCB1_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Mortimore B.;
RL Submitted (Oct-1996) to the EMBL/Genbank/DBS databases.
CC -1- CATALYTIC ACTIVITY: 2 glutathione + H(2)O(2) = oxidized
CC glutathione + 2 H(2)O.
CC -1- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC -1- SIMILARITY: Belongs to the glutathione peroxidase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z81015; CAB02655.1; -.
DR PIR; T19190; T19190.
DR HSSP; P00435; 1GPI.
DR WormPep; C1B4.2; CB08102.
DR InterPro; IPR000889; Glut_peroxidase.
DR Pfam; PF00255; GSHPX; 1.
DR PRINTS; PR01011; GLUTATHIONE_PEROXIDASE.
DR PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.

```


DR PROSITE; PS00763; GLUTATHIONE PEROXIDASE 2; 1.
 KM Oxidoreductase; Peroxidase; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 224
 FT ACT SITE 73 73
 FT CAROTHD 38 38
 FT SEQUENCE 224 AA; 25556 MW; F02D055246DE2F1 CRC64;
 SQ
 Query Match
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTD 5
 Db 77 QOYTD 81
 RESULT 10
 STX7 HUMAN STANDARD; PRT; 261 AA.
 015400; Q96E89;
 15-JUN-1998 (Rel. 36, Created)
 28-FEB-2003 (Rel. 41, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Syntaxin 7.
 GN STX7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OK NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=98019069; PubMed=9358037;
 RA Wang H., Prelin L., Pevenner J.,
 "Human syntaxin 7: a Pep12p/Vps6p homologue implicated in vesicle
 trafficking to lysosomes.";
 RT Gene 199:39-48(1997).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE-Ovary;
 RC MEDLINE=22388257; PubMed=12477932;
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ueidi T.B., Toshiyuki S., Carninci P., Prange C.,
 Rana S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalon D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butlerfield Y.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra W.A.,
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=20032182; PubMed=10564279;
 RA Prekeris R., Yang B., Oorschot V., Klumperman J., Scheller R.H.,
 "Differential roles of syntaxin 7 and syntaxin 8 in endosomal
 trafficking.";
 RT Mol. Biol. Cell 10:3891-3908(1999).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=20158968; PubMed=10692457;
 RA Nakamura N., Yamamoto A., Wada Y., Futai M.,

RT "Syntaxin 7 mediates endocytic trafficking to late endosomes.";
 RL J. Biol. Chem. 275:6523-6529(2000).
 CC -1- FUNCTION: May be involved in protein trafficking from the plasma
 membrane to the early endosome (EE) as well as in homotypic fusion
 of endocytic organelles. Mediates the endocytic trafficking from
 early endosomes to late endosomes and lysosomes.
 CC -1- SUBUNIT: Forms a snare complex with Vti1b, STX8 AND
 endobrevin/VAMP8 which functions in the homotypic fusion of late
 endosomes. Interacts with VPS33A (By similarity). Interacts with
 VPS11, VPS16 and VPS18.
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein. Early endosome
 membranes (By similarity).
 CC -1- TISSUE SPECIFICITY: Highest expression is found in placenta
 followed by heart, skeletal muscle, kidney and liver.
 CC expression is found in pancreas, lung and liver.
 CC -1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.
 CC -1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U77942; AAC51851.1; -
 DR EMBL; BC011975; AAH11975.1; -
 DR Genew; HGNC:11442; STX7.
 DR MIM; 603217; -
 DR GO; GO:0006892; P:post Golgi transport; TAS.
 DR InterPro; IPR006012; Syntaxin.
 DR InterPro; IPR006011; Syntaxin_N.
 DR InterPro; IPR000727; T-SNARE.
 DR Pfam; PF00804; Syntaxin; 1.
 DR SMART; SM00503; SYN; 1.
 DR SMART; SM00397; t-SNARE; 1.
 DR PROSITE; PS00914; SYNTAXIN; 1.
 DR PROSITE; PS50192; T-SNARE; 1.
 DR Coiled coil; Transmembrane.
 FT FT
 FT DOMAIN 1 258
 FT TRANSMEM 239 259
 FT FT
 FT DOMAIN 260 261
 FT DOMAIN 47 69
 FT DOMAIN 165 227
 FT CONFLICT 12 12
 FT SEQUENCE 261 AA; 29815 MW; 42AC173F0233ACDA CRC64;
 SQ
 Query Match
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTD 6
 Db 56 QOYTD 61
 RESULT 11
 STX7 MOUSE STANDARD; PRT; 261 AA.
 ID STX7 MOUSE
 AC 070439;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Syntaxin 7.
 GN STX7 OR SYN7.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Adipocyte;
 RA Tellam J., Piper R.C., Smith C., James D.B.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: May be involved in protein trafficking from the plasma
 CC membrane to the early endosome (EE) as well as in homotypic fusion
 CC of endocytic organelles. Mediates the endocytic trafficking from
 CC early endosomes to late endosomes and lysosomes (By similarity).
 CC -1- SUBUNIT: Forms a snare complex with VTI1B, STX8 AND
 CC endobrevin/VAMP8 which functions in the homotypic fusion of late
 CC endosomes. Interacts with VPS11, VPS16, VPS18 and VPS33A (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein. Early endosome
 CC membranes (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.
 CC -1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF056323; AAC15971.1; -;
 CC MGD; MG1:1858210; Str7.
 DR InterPro; IPR006012; Syntaxin.
 DR InterPro; IPR006011; Syntaxin_N.
 DR InterPro; IPR000727; T_SNARE.
 DR Pfam; PF00804; Syntaxin; 1.
 DR SMART; SM00503; SYN; 1.
 DR SMART; SM00397; t-SNARE; 1.
 DR PROSITE; PS00914; SYNTAXIN; 1.
 DR PROSITE; PS50192; T_SNARE; 1.
 KW Coiled coil; Transmembrane.
 DR DOMAIN 1 238 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 239 259 ANCHOR FOR TYPE IV MEMBRANE PROTEIN
 FT DOMAIN 260 261 (POTENTIAL).
 FT DOMAIN 47 68 ENDOSOMAL (POTENTIAL).
 FT DOMAIN 165 227 COILED COIL (POTENTIAL).
 FT DOMAIN 227 227 T-SNARE COILED-COIL HOMOMOLOGY.
 SQ SEQUENCE 261 AA; 29820 MW; 62F4E4B33F247ADB CRC64;
 Query Match 84.8%; Score 28; DB 1; Length 261;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 QOYTDQ 6
 |||||
 56 QOYTNQ 61
 RESULT 12
 STX7 RAT STANDARD; PRT; 261 AA.
 ID STX7 RAT
 AC 070257;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Syntaxin 7.
 GN STX7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 OX (1)
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=98079071; PubMed=9417091;
 RT Wong S.H., Xu Y., Zhang T., Hong W.;
 RT "Syntaxin 7, a novel syntaxin member associated with the early
 RT endosomal compartment.";
 RL J. Biol. Chem. 273:375-380(1998).

RN [2]
 RP SEQUENCE FROM N.A., AND SNARE COMPLEX CHARACTERIZATION.
 RX MEDLINE=20553225; PubMed=11101518.
 RA Antonin W., Holroyd C., Faehauer D., Pabst S.,
 RA Fischer von Mollard G., Jahn R.;
 RT "A SNARE complex mediating fusion of late endosomes defines conserved
 RT properties of SNARE structure and function.";
 RL EMBL J. 19:6453-6464(2000)
 CC -1- FUNCTION: May be involved in protein trafficking from the plasma
 CC membrane to the early endosome (EE) as well as in homotypic fusion
 CC of endocytic organelles. Mediates the endocytic trafficking from
 CC early endosomes to late endosomes and lysosomes (By similarity).
 CC -1- SUBUNIT: Forms a snare complex with VTI1B, STX8 AND
 CC endobrevin/VAMP8 which functions in the homotypic fusion of late
 CC endosomes. Interacts with VPS11, VPS16 and VPS18 (By similarity).
 CC interacts with VPS33A.
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein. Early endosome
 CC membranes (By similarity).
 CC -1- TISSUE SPECIFICITY: Detected in all tissues tested. Highest
 CC expression is found in kidney followed by lung, spleen, heart and
 CC brain. Lower expression, in skeletal muscle, liver and testis.
 CC -1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.
 CC -1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF031430; AAC17131.1; -;
 DR InterPro; IPR006012; Syntaxin.
 DR InterPro; IPR006011; Syntaxin_N.
 DR InterPro; IPR000727; T_SNARE.
 DR Pfam; PF00804; Syntaxin; 1.
 DR SMART; SM00503; SYN; 1.
 DR SMART; SM00397; t-SNARE; 1.
 DR PROSITE; PS00914; SYNTAXIN; 1.
 DR PROSITE; PS50192; T_SNARE; 1.
 KW Coiled coil; Transmembrane.
 DR DOMAIN 1 238 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 239 259 ANCHOR FOR TYPE IV MEMBRANE PROTEIN
 FT DOMAIN 260 261 (POTENTIAL).
 FT DOMAIN 47 68 ENDOSOMAL (POTENTIAL).
 FT DOMAIN 165 227 COILED COIL (POTENTIAL).
 FT CONFLICT 139 139 T-SNARE COILED-COIL HOMOMOLOGY.
 FT CONFLICT 139 139 A -> E (IN REF. 2).
 SQ SEQUENCE 261 AA; 29792 MW; 44123967B85F5F09 CRC64;
 Query Match 84.8%; Score 28; DB 1; Length 261;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 QOYTDQ 6
 |||||
 56 QOYTNQ 61
 Db
 YG1T YEAST STANDARD; PRT; 271 AA.
 ID YG1T YEAST
 AC P53227;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 31.0 kDa protein in BUD9-RME1 intergenic region.
 GN YG042W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 Kilobases from Saccharomyces cerevisiae
   chromosome VII.";
RL Yeast 13:1077-1090(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Entian K.D., Rose M., Koetter P., Roehmer A., Sehraam I.,
   Hempel S.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 272827; CAA97040.1; -
DR PIR; S64333; S64333.
DR SGD; S0003274; YGR042M.
KM Hypothetical protein.
SQ SEQUENCE 271 AA; 31025 MW; 5208A41734A31520 CRC64;

Query Match      84.8%; Score 28; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OQYTD 6
DB 11 OQYTD 15

RESULT 14
META YERPE
ID _META_YERPE STANDARD; PRT; 309 AA.
AC _O8ZAR4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homoserine O-succinyltransferase (EC 2.3.1.46) (Homoserine O-
   transsuccinylase) (HRS).
GN METR OR YP03727 OR Y0013.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
   Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parthill J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
   Prentice M.B., Sepahinia M., James K.D., Churcher C., Mungall K.L.,
   Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
   Feltham D., Hamlin N., Holtroyd S., Jagsis K., Karlyshev A.V.,
   Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RA "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G., III, Boutin A., Mayhew G.F., Liss P.,
   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
   Fetherston J.D., Lindler L.B., Brubaker R.R., Plano G.V.,
   Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
   Perry R.D.;

```

```

RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
CC -I- CATALYTIC ACTIVITY: Succinyl-CoA + L-homoserine = CoA + O-
   succinyl-L-homoserine.
CC -I- PATHWAY: Methionine biosynthesis; HTS variant; first step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: BELONGS TO THE HTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ414158; CAC93195.1; -
DR EMBL; AE013602; AAM83609.1; -
DR PIR; AG0453; AG0453.
DR HAMAP; MF_00295; -; 1.
DR InterPro; IPR005697; Homoser_succ.
DR Pfam; PF04204; HRS; 1.
DR TIGRPFAM; TIGR01001; metA; 1.
KM Methionine biosynthesis; Transferase; Acyltransferase;
   Complete proteome.
FT ACT SITE 142 142
SQ SEQUENCE 309 AA; 35365 MW; 1349B9999514EAB8 CRC64;

Query Match      84.8%; Score 28; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OQYTD 5
DB 203 OQYTD 207

RESULT 15
NHR7_CABEL
ID _NHR7_CABEL STANDARD; PRT; 536 AA.
AC Q20765;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Nuclear hormone receptor family member nhr-7.
GN NHR-7 OR E54D1.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
   Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99148134; PubMed=10022975;
RA Sluder A.E., Mathews S.W., Hough D., Yin V.P., Maina C.V.;
RT "The nuclear receptor superfamily has undergone extensive
   proliferation and diversification in nematodes.";
RL Genome Res. 9:103-120(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Leonard N.;
RL Submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Durbin R.;
RL Submitted (Jan-2002) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: Orphan nuclear receptor.
CC -I- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -I- SIMILARITY: Belongs to the nuclear hormone receptor family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)

DR EMBL; AF083325; AAD03683.1; -
 DR EMBL; 277133; CAB00862.3; -
 DR PIR; T43349; T43349.
 DR HSSP; P19793; 2NLT.
 DR WormPep; F54D1.4; CE20866.
 DR InterPro; IPR000536; Hormone_rec_119.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STROIDFINGER.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 Zinc-finger.

FT DNA BIND 9 75 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 9 29 C4-TYPE.
 FT ZN_FING 46 70 C4-TYPE.
 SQ SEQUENCE 536 AA; 60424 MW; 07B95116D41A378 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 536;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 00QYTD 5
 Db 439 00QYTD 443

RESULT 16
 C3BB_BACTU STANDARD; PRT; 652 AA.
 ID C3BB_BACTU 006117; 045717;
 AC 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Pesticidal crystal protein cry3Bb (insecticidal delta-endotoxin
 DE CryIIIBb) (Crystalline entomocidal protoxin) (74 kDa crystal
 DE protein).
 GN CRY3BB OR CRYIIIB(B) OR CRYIIIB2.
 OS Bacillus thuringiensis.
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=1428;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EG4961;
 RX MEDLINE=93119147; PubMed=1476436;
 RA Donovan W.P., Ruper M.J., Stanley A.C., Malvar T., Gawron-Burke M.C.,
 RA Johnson T.B.;
 RT "Characterization of two genes encoding Bacillus thuringiensis
 RT insecticidal crystal proteins toxic to Coleoptera species.";
 RT Appl. Environ. Microbiol. 58:3921-3927(1992).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=NRRL B-18655 / EG5144;
 RX Donovan W.P., Ruper M.J., Stanley A.C.;
 RA "Bacillus thuringiensis cryIIIC, (b) protein toxic to coleopteran
 RA insects.";
 RT Patent number US5378625, 03-JAN-1995.
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF COLEOPTERA. HAS MODERATE LEVEL OF TOXICITY TO
 CC SOUTHERN CORN ROOTWORM.
 CC -1- SUBUNIT: Monomer.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART OF
 CC THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE N-

CC TERMINUS. BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)

DR EMBL; M89794; AAA22334.1; -
 DR EMBL; U31633; AAA74198.1; -
 DR PIR; I39811; I39811.
 DR PDB; 1J16; 19-SEP-01.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; Endotoxin_C.
 DR InterPro; IPR005639; Endotoxin_N.
 DR Pfam; PF00555; endotoxin_1.
 DR Pfam; PF03944; endotoxin_C; 1.
 DR Pfam; PF03945; endotoxin_N; 1.
 KM Toxin; Sporulation; 3D-structure.
 FT VARIANT 21 21 Q -> P (IN STRAIN EG5144).
 FT VARIANT 97 97 N -> D (IN STRAIN EG5144).
 FT VARIANT 289 289 I -> V (IN STRAIN EG5144).
 FT VARIANT 352 352 F -> S (IN STRAIN EG5144).
 FT VARIANT 417 419 VYL -> IYF (IN STRAIN EG5144).
 FT VARIANT 451 451 S -> G (IN STRAIN EG5144).
 FT VARIANT 590 590 L -> I (IN STRAIN EG5144).
 FT VARIANT 600 600 K -> I (IN STRAIN EG5144).
 FT VARIANT 624 624 K -> T (IN STRAIN EG5144).
 SQ SEQUENCE 652 AA; 74385 MW; 63048332CDB8CC6 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 652;
 Best Local Similarity 100.0%; Pred. No. 116+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 00QYTD 5
 Db 238 00QYTD 242

RESULT 17
 C3BA_BACTO STANDARD; PRT; 659 AA.
 ID C3BA_BACTO P17969;
 AC 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Pesticidal crystal protein cry3Ba (insecticidal delta-endotoxin
 DE CryIIIBa) (Crystalline entomocidal protoxin) (75 kDa crystal
 DE protein).
 GN CRY3BA OR CRYIIIB(A) OR CRYIIIB.
 OS Bacillus thuringiensis (subsp. tolworthi).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=1442;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=43F;
 RX MEDLINE=90206811; PubMed=2320431;
 RA Sick A., Gaertner F.H., Wong A.;
 RT "Nucleotide sequence of a coleopteran-active toxin gene from a new
 RT isolate of Bacillus thuringiensis subsp. tolworthi.";
 RT Nucleic Acids Res. 18:1305-1305(1990).
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF COLEOPTERA.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X17123; CAA34983.1; -
 DR EMBL; A07234; CAA00645.1; -
 DR PIR; S10228; S10228.
 DR HSSP; P07130; IDIC.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; endotoxin_C.
 DR InterPro; IPR005639; endotoxin_N.
 DR Pfam; PF00555; endotoxin_1.
 DR Pfam; PF03944; endotoxin_C; 1.
 DR Pfam; PF03945; endotoxin_N; 1.
 DR Toxin; Sporulation.
 KW SEQUENCE 659 AA; 75159 MW; 5A5B214FF84168CA CRC64;

Query Match 84.8%; Score 28; DB 1; Length 659;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTD 5
 DB 246 QOYTD 250

RESULT 18
 ID R161 YEAST STANDARD; PRT; 265 AA.
 AC P25343;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Reduced viability upon starvation protein 161.
 GN RVS161 OR SPI61 OR BND6 OR YCR090C OR YCR9C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A. / SKO2N;
 RC STRAIN=ATCC-44827;
 RX MEDLINE=90224366; PubMed=2183524;
 RT Urdaci M., Dulau L., Aigle M., Crouzet M.;
 RN "Sequence of the yeast gene RVS 161 located on chromosome III.";
 RN Yeast 6:173-176(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=92327849; PubMed=1626432;
 RX Skala J., Purnelle B., Goffeau A.;
 RN "The complete sequence of a 10.8 kb segment distal of SUP2 on the
 RT right arm of chromosome III from Saccharomyces cerevisiae reveals
 RT seven open reading frames including the RVS161, ADP1 and PGK genes.";
 RN Yeast 8:409-417(1992).
 RN [3]
 RP CHARACTERIZATION.
 RC STRAIN=ATCC 44827 / SKO2N;
 RX MEDLINE=92133163; PubMed=1776363;
 RT Crouzet M., Urdaci M., Dulau L., Aigle M.;
 RN "Yeast mutant affected for viability upon nutrient starvation:
 RT characterization and cloning of the RVS161 gene.";
 RN Yeast 7:727-743(1991).
 CC -1- FUNCTION: COMPONENT OF A CYTOSKELETAL STRUCTURE THAT IS REQUIRED
 CC FOR THE FORMATION OF ENDOCYTIC VESICLES AT THE PLASMA MEMBRANE
 CC LEVEL.
 CC -1- MISCELLANEOUS: MUTATIONS IN THIS GENE RESULTS IN SENSITIVITY TO
 CC CARBON, NITROGEN AND SULFUR STARVATION.
 CC -1- SIMILARITY: Contains 1 BAR domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X59720; CAA42326.1; -
 DR EMBL; X63315; CAA44926.1; -
 DR PIR; S19516; S19516.
 DR SGD; S0000602; RVS161.
 DR GO; GO:0005857; C:actin cortical patch (sensu Saccharomyces); IDA.
 DR GO; GO:0045121; C:lipid raft; IDA.
 DR GO; GO:0008092; F:cytoskeletal protein binding activity; IPI.
 DR GO; GO:0006897; P:endocytosis; IMP.
 DR GO; GO:0007121; P:polar budding; IMP.
 DR GO; GO:0006970; P:response to osmotic stress; IMP.
 DR InterPro; IPR006632; BAR.
 DR InterPro; IPR004148; BAR_dom.
 DR Pfam; PF03114; BAR; 1.
 DR SMART; SM00721; BAR; 1.
 KW Cytoskeleton; Coiled coil.
 FT DOMAIN 24 56
 FT DOMAIN 126 193
 FT CONFLICT 95 95
 FT CONFLICT 95 95
 FT O -> E (IN REF. 1).
 SQ SEQUENCE 265 AA; 30250 MW; B35EA00B0C900E2D CRC64;

Query Match 81.8%; Score 27; DB 1; Length 265;
 Best Local Similarity 83.3%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTD 6
 DB 227 QOYTD 232

RESULT 19
 ID HEMZ PASMU STANDARD; PRT; 326 AA.
 AC P57874;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ferrochelatase (EC 4.99.1.1) (Prothome ferro-lyase) (Heme
 DE synthetase).
 GN HEMZ OR PM0789.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxId=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RX MEDLINE=21145866; PubMed=11248100;
 RT May B.J., Zhang O., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RN "Complete genomic sequence of Pasteurella multocida Pm70.";
 RN Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 RN [2]
 RP FUNCTION: CATALYZES THE FERROUS INSERTION INTO PROTOPORPHYRIN IX.
 CC -1- CATALYTIC ACTIVITY: Protoporphyrin + Fe(2+) = prothome + 2 H(+).
 CC -1- PATHWAY: Prothome biosynthesis; last step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the ferrochelatase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB006117; AAK02873.1; -

DR HAMAP; MF 00323; -; 1.
 DR InterPro: IPR001015; Ferrochelatase.
 DR Pfam: PF00762; Ferrochelatase; 1.
 DR ProDom: PD002792; Ferrochelatase; 1.
 DR TrRfam: TRF00109; hemh; 1.
 DR PROSITE; PS00534; FERROCHELATASE; 1.
 DR Porphyria biosynthesis; Heme biosynthesis; Lyase; Iron;
 KM Complete proteome.
 FT METAL 195 195 IRON (BY SIMILARITY).
 FT METAL 276 276 IRON (BY SIMILARITY).
 SQ SEQUENCE 326 AA; 37421 MW; 5AAAEAC7DE3D5 CRC64;
 Query Match 81.8%; Score 27; DB 1; Length 326;
 Best Local Similarity 83.3%; Pred. No. 86;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 DB 245 QPYTDQ 250
 RESULT 20
 ID CV05_HUMAN STANDARD; PRT; 373 AA.
 AC Q9Y519; Q9UGT8; Q9UGT9; Q9UGV5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein C22orf5.
 GN C22ORF5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Collins J.E., Huckley E.J.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clamp M., Smit L.J., Alnecough R., Almeida J.P., Babbage A.K.,
 RA Baggaley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carter C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt R.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leverisa M.A., Lloyd C., Lloyd D.M.,
 RA Marley I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
 RA McClellan J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Senra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spegren L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kodoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mituyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren O., Shauli S., Sloan D., Song L.,
 RA Zhan Q., Wang Y., Wang Z., White J., Williamson D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Mink P.,
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goele D., Graves T., Hawkins J.,

RA Hinde K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Woldmann P., Pepin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
 RA Emanuel B.S., Shalko T., Kurahashi H., Salta S., Budarf M.L.,
 RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
 RA Kim U.J., Shizuya H., Simon M.I., Dunneki J.P., Payard M., Kedra D.,
 RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
 RA Wilkerson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
 RA Tikhonov Y., Wright H.;
 RT "The DNA sequence of human chromosome 22."
 RL Nature 402:489-495 (1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Ueda T.B., Toshiyuki S., Cantucci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.D., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: BELONGS TO THE UPF0206 FAMILY.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 gene model prediction.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 DR EMBL; AL096879; CAB51403.1; -;
 DR EMBL; AL020993; CAB63065.1; ALT_SEQ.
 DR EMBL; AL020993; CAB63066.1; ALT_SEQ.
 DR EMBL; AL021977; CAB45267.1; -;
 DR EMBL; BC015489; AAH15489.1; -;
 DR EMBL; BC015489; AAH15489.1; -;
 DR EMBL; BC015489; AAH15489.1; -;
 DR InterPro: IPR005178; DUF300.
 DR Pfam: PF03619; DUF300; 1.
 KW Transmembrane.
 FT TRANSMEM 6 26 POTENTIAL.
 FT TRANSMEM 50 70 POTENTIAL.
 FT TRANSMEM 87 107 POTENTIAL.
 FT TRANSMEM 144 164 POTENTIAL.
 FT TRANSMEM 180 200 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 256 276 POTENTIAL.
 SQ SEQUENCE 373 AA; 42329 MW; 28C4CEB220A4EAF CRC64;
 Query Match 81.8%; Score 27; DB 1; Length 373;
 Best Local Similarity 83.3%; Pred. No. 99;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 DB 328 QOYTDQ 333

RESULT 21
 ID SVE STRP3 STANDARD; PRT; 481 AA.
 AC ORKBP9; 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
 DE (Glurs)
 GN GLTX OR GLUS OR SPYM_0170 OR SPS0177.
 OS Streptococcus pyogenes (serotype M3).
 OC Bacteria; Firmicutes; Lactobactillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=198466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGAS315 / Serotype M3;
 MEDLINE=22133808; PubMed=12122206;
 Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 Schlievert P.M., Musser J.M.;
 "Genome sequence of a serotype M3 strain of group A Streptococcus:
 phase-encoded toxins, the high-virulence phenotype, and clone
 emergence.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SSI-1 / Serotype M3;
 RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
 Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
 Hayashi H., Hamada S.;
 "The genome of invasive Streptococcus pyogenes; a comparative analysis
 of S. pyogenes SSI-1, SF370 and MGAS832.";
 RT Submitted (MAA-2002) to the EMBL/GenBank/DBJ databases.
 RL -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
 -1- dihydrophosphate + L-glutamyl-tRNA(Glu).
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AE014139; AAM78777.1; -;
 DR EMBL, AP005141; BAC63272.1; -;
 DR HAMAP, MF_00022; -; 1.
 DR InterPro; IPR004527; GLX Bact.
 DR InterPro; IPR000924; Glu tRNA-synt_1c.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR Pfam; PF00749; tRNA-synt_1c; 1.
 DR PRINTS; PR00987; TRNASYNTHGU.
 DR TIGRFAMs; TIGR00464; glx Bact; 1.
 DR PROSITE; PS00178; AA tRNA_LIGASE I; 1.
 DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 DR SITE 11 21 "HIGH" REGION.
 FT SITE 255 259 "KMSKS" REGION.
 FT BINDING 258 258 ATP (BY SIMILARITY).
 SQ SEQUENCE 481 AA; 55063 MW; D03586D0CD92D082 CRC64;

Query Match 81.8%; Score 27; DB 1; Length 481;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 OQYTDQ 6

Db 90 OQYTDQ 95
 RESULT 22
 ID SVE STRPY STANDARD; PRT; 481 AA.
 AC O9AIJ8; 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
 DE (Glurs)
 GN GLTX OR GLUS OR SPY0239 OR SPYM18_0223.
 OS Streptococcus pyogenes, and
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobactillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314, 186103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 MEDLINE=21192684; PubMed=11296296;
 Ferretti J.J., Moshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 Primeaux C., Sezesh S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 Qian Y., Jia H.G., Najaf F.Z., Ren Q., Zhu H., Song L., White J.,
 Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGAS832 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108;
 Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 Sylva G.L., Sturdevant D.E., Ricklets S.M., Porcella S.F.,
 Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 "Genome sequence and comparative microarray analysis of serotype M18
 group A Streptococcus strains associated with acute rheumatic fever
 outbreaks.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 RL -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
 -1- dihydrophosphate + L-glutamyl-tRNA(Glu).
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AE006491; AAK33319.1; -;
 DR EMBL, AE009971; AAL37010.1; -;
 DR HSSP; P27000; IGLN.
 DR HAMAP, MF_00022; -; 1.
 DR InterPro; IPR004527; GLX Bact.
 DR InterPro; IPR000924; Glu tRNA-synt_1c.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR Pfam; PF00749; tRNA-synt_1c; 1.
 DR PRINTS; PR00987; TRNASYNTHGU.
 DR TIGRFAMs; TIGR00464; glx Bact; 1.
 DR PROSITE; PS00178; AA tRNA_LIGASE I; 1.
 DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 DR SITE 11 21 "HIGH" REGION.
 FT SITE 255 259 "KMSKS" REGION.
 FT BINDING 258 258 ATP (BY SIMILARITY).
 SQ SEQUENCE 481 AA; 55093 MW; 984087D5C0EAA08B CRC64;

Query Match 81.8%; Score 27; DB 1; Length 481;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 90 QOYTDQ 95

RESULT 23

OSTA_PASMU STANDARD; PRT; 782 AA.
 AC O9CKL2;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ORGANIC solvent tolerance protein precursor.
 GN IMP OR OSTA OR PM1600.
 OS Pasteurella multocida.
 NC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 NC Pasteurellaceae; Pasteurella.
 NC NCBI_TaxID=747;

RE SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang O., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.,
 RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 RL "Complete genomic sequence of Pasteurella multocida PM70.";
 CC -1- FUNCTION: Determines N-hexane tolerance. Involved in outer
 CC membrane permeability. Essential for envelope biogenesis. Could be
 CC part of a targeting/usher system for outer membrane components (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Outer membrane (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE IMP/OSTA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE006197; AAK03684.1; ALT_INIT.

DR HAMAP: MF_01411; -; 1.

DR InterPro: IPR005653; Osta.

DR Pfam: PF03968; Osta_C; 1.

DR Pfam: PF04453; Osta_C; 1.

DR Outer membrane; Signal; Complete proteome.

FT SIGNAL 1 24 POTENTIAL.

FT CHAIN 25 782 ORGANIC SOLVENT TOLERANCE PROTEIN.

FT SEQUENCE 782 AA; 90647 MW; 557961B12B2031E6 CRC64;

QY 1 QOYTDQ 6

DB 657 QOYTDQ 662

RESULT 24

ATRX_MOUSE STANDARD; PRT; 2476 AA.

AC Q61687;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-SRP-2003 (Rel. 42, Last annotation update)

DE Transcriptional regulator ATRX (X-linked nuclear protein)

DE (Heterochromatin protein 2) (HPI alpha-interacting protein) (HPI-BP38
 DE protein).

GN ATRX OR XNP OR HPIBP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.
 RX MEDLINE=98213653; PubMed=9545503;
 RA Picketts D.J., Taetan A.O., Higgs D.R., Gibbons R.J.,
 RT "Comparison of the human and murine ATRX gene identifies highly
 RT conserved, functionally important domains."
 RL Mamm. Genome 9:400-403(1998).

RP SEQUENCE OF 325-1176 FROM N.A.
 RX MEDLINE=97133299; PubMed=8978696;
 RA le Douarin B., Nielsen A.L., Garnier J.-M., Ichinose H.,
 RA Jeannouin F., Loison R., Chambon P.,
 RT "A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic
 RT control of transcription by nuclear receptors."
 RL EMBO J. 15:6701-6715(1996).

RP [3]
 RP SUBCELLULAR LOCATION: AND ASSOCIATION WITH PERICENTROMERIC
 RP HETEROCHROMATIN.

RX MEDLINE=20040663; PubMed=10570185;
 RA McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,
 RA Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,
 RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.,
 RT "Localisation of a putative transcriptional regulator (ATRX) at
 RT pericentromeric heterochromatin and the short arms of acrocentric
 RT chromosomes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).

CC -1- FUNCTION: COULD BE A GLOBAL TRANSCRIPTIONAL REGULATOR. MODIFIES
 CC GENE EXPRESSION BY AFFECTING CHROMATIN.
 CC -1- SUBUNIT: PROBABLY BINDS EZH2. BINDS ANNEKIN V IN A CALCIUM AND
 CC PHOSPHATIDYLCHOLINE/PHOSPHATIDYLSELINE-DEPENDENT MANNER (By
 CC similarity).

CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH PERICENTROMERIC
 CC HETEROCHROMATIN DURING INTERPHASE AND MITOSIS, PROBABLY BY
 CC INTERACTING WITH HPI.
 CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.

CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF026032; AAC08741.1;
 DR EMBL: X99643; CA67962.1; -;
 DR GGD; GGD:103067; Atrx.

DR GO: GO:0000228; C:nuclear chromosome; IDA.

DR InterPro: IPR001410; DEAD.

DR InterPro: IPR001650; Helicase_C.

DR InterPro: IPR000330; SNF2_N.

DR InterPro: IPR001841; Znf_finger.

DR Pfam: PF00271; helicase_C; 1.

DR Pfam: PF00176; SNF2_N; 1.

DR SMART; SM00487; DEXDC; 1.

KW DNA repair; Nuclear protein; DNA-binding; Helicase; ATP-binding;
 KW Zinc-finger.

FT ZN_FINGER 219 267 PHD-TYPE.

FT NP_BIND 1579 1586 ATP (POTENTIAL).

FT SITE 1704 1707 DEGR_BOX.

FT DOMAIN 319 322 POLY-SER.

FT DOMAIN 735 738 POLY-SER.

FT DOMAIN 1001 1004 POLY-GLU.

FT DOMAIN 1130 1135 POLY-SER.

FT DOMAIN 1182 1185 POLY-SER.

FT DOMAIN 1238 1245 POLY-ASP.

FT DOMAIN 1484 1487 POLY-GLU.


```

FT DOMAIN 1924 1931 POLY-SER.
FT DOMAIN 2205 2208 POLY-LYS.
FT DOMAIN 2245 2248 POLY-GLU.
FT DOMAIN 2403 2408 POLY-GLN.
SQ SEQUENCE 2476 AA; 278601 MW; 90A42B790FC4FFAC CRC64;

Query Match
Best Local Similarity 81.8%; Score 27; DB 1; Length 2476;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 OQYTDQ 6
DB 2399 OQYTDQ 2404

RESULT 25
YIID ECOLI STANDARD; PRT; 329 AA.
AC P32148;
01-OCT-1993 (Rel. 27, Created)
01-OCT-1993 (Rel. 27, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Hypothetical protein yiid.
DE YIID OR B3888 OR C4835 OR Z5427 OR ECG4811.
OS Escherichia coli,
OS Escherichia coli O6, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562, 217992, 83334;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RC MEDLINE=93347969; PubMed=8346018;
RX Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
RA "Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes."
RL Nucleic Acids Res. 21:3391-3398(1993).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=O6:H1 / CFT073 / ATCC 700928;
RC MEDLINE=2288234; PubMed=12471157;
RX Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RA "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Poserai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Poromousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kubera S., Shiba T., Hattori M., Shingawa H.;
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
CC

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19201; AAB03021.1; -
DR EMBL; AE009464; AAD13450.1; -
DR EMBL; AE016770; AAN83264.1; -
DR EMBL; AE005619; AAG59078.1; -
DR EMBL; AP002567; BAB38234.1; -
DR PIR; C91230; C91230.
DR PIR; S40832; S40832.
DR Ecogene; Bg11853; yiid.
DR InterPro; IPR00182; GCN5acetyltransf.
DR Pfam; PF00583; Acetyltransf; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 329 AA; 37094 MW; 42411FE2B0D21175 CRC64;

Query Match
Best Local Similarity 78.8%; Score 26; DB 1; Length 329;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 OQYTDQ 6
DB 196 OQYTDQ 201

RESULT 26
DNI2 CAEBL STANDARD; PRT; 337 AA.
ID DNI2 CAEBL
AC 017433;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DR DnaJ homolog dnf-2 precursor.
GN DNF-2 OR B0035.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodertinae; Caenorhabditis.
OX NCBI_TaxId=6239;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol NZ;
RA White S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 J domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73102; CAA97409.1; -
DR PIR; T18654; T18654.
DR WormPep; B0035.2; CE05161.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ; 1.
DR SMART; SM00227; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
KW Signal; Chaperone.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 337 DnaJ HOMOLOG DNF-2.
FT DOMAIN 36 105 J DOMAIN.
SQ SEQUENCE 337 AA; 39992 MW; 03CB5511FE0185D CRC64;

Query Match
Best Local Similarity 78.8%; Score 26; DB 1; Length 337;

```

Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 276 QOYTD 281

RESULT 27

GBA2_CAEEL STANDARD; PRT; 356 AA.
AC P22454;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Guanine nucleotide-binding protein alpha-2 subunit.
GN CPA-2 OR F38E1.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
NCBI_TaxID=6239;

SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=91039310; PubMed=2121996;
RA Silva F.I., Plasterk R.H.A.;
RT "Characterization of a G-protein alpha-subunit gene from the nematode
RT Caenorhabditis elegans."
RT J. Mol. Biol. 215:483-487(1990).

SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Cuppen E., Jansen G., Plasterk R.H.A.;
RT "Interaction analysis of the complete G-alpha subfamily of
RT heterotrimeric G proteins from Caenorhabditis elegans."
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Gattung S., Le T.T.;
RT Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEIN (G PROTEINS) ARE
INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
SIGNALING SYSTEMS.

-1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS: ALPHA, BETA AND
GAMMA. THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING
SITE.

-1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X53156; CA37312.1; -
EMBL; AY008125; AAG32078.1; -
EMBL; U41996; AAA83475.1; -
DR PIR; S13221; S13221.
DR HSSP; P04896; 1AZT.
DR WormBep; F38E1.5; CE04523.
DR InterPro; IPR001019; Gprotelin_alpha.
DR Pfam; PF00503; G-alpha_1.
DR PRINTS; PR00318; GPROTEIN.
DR ProDom; PD000281; Gprotelin_alpha; 1.
DR SMART; SM00275; G_alpha; 1.
DR GTP-binding; Transducer; Multigene family.
FT NP_BIND 40 47 GTP (BY SIMILARITY).
FT NP_BIND 203 207 GTP (BY SIMILARITY).
FT NP_BIND 272 275 GTP (BY SIMILARITY).
SEQUENCE 356 AA; 41489 MW; B2622850B3AF04ED CRC64;

Query Match 78.8%; Score 26; DB 1; Length 356;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 59 QOYTD 64

RESULT 28

COB8_SYNY3 STANDARD; PRT; 482 AA.
AC P73002;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cobytrinic acid A.C-diamide synthase.
GN COB8 OR CBIA OR SLL1501.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
NCBI_TaxID=1148;

SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsumoto A., Muraki A., Nakazaki N., Maruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RT DNA Res. 3:109-116(1996).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D90902; BAA17022.1; -
DR PIR; S74982; S74982.
DR HAMAP; MF 00027; -; 1.
DR InterPro; IPR004484; CBIA.
DR InterPro; IPR002586; CBIA_P.
DR Pfam; PF01656; CBIA_1.
DR TIGRFAMs; TIGR00379; COB8; 1.
KW Cobalamin biosynthesis; Porphyrin biosynthesis; Complete proteome.
SQ SEQUENCE 482 AA; 53854 MW; FA0A5C17B9C4408F CRC64;

Query Match 78.8%; Score 26; DB 1; Length 482;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 202 QOYTDQ 207

RESULT 29

CPH1_CANAL STANDARD; PRT; 656 AA.
ID CPH1_CANAL
AC P43079;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcription factor CPH1.
GN CPH1 OR ACPR.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxId=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SC5314;
 RX MEDLINE=94364983; PubMed=8083193;
 RA Malachuk K., Ganesan K., Datta A.;
 RT "Identification of a putative transcription factor in Candida
 albicans that can complement the mating defect of Saccharomyces
 cerevisiae str12 mutants";
 RL J. Biol. Chem. 269:22945-22951(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95084165; PubMed=7992058;
 RA Liu H., Koehler J., Fink G.R.;
 RT "Suppression of hyphal formation in Candida albicans by mutation of a
 STE12 homolog";
 RL Science 266:1723-1726(1994).
 RN [3]
 RP ERRATUM.
 RA Liu H., Koehler J., Fink G.R.;
 RL Science 267:17-17(1994).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=95100929; PubMed=7802635;
 RA Singh P., Ganesan K., Malachuk K., Ghosh D., Datta A.;
 RT "ACR, a STE12 homologue from Candida albicans, is a strong inducer
 of pseudohyphae in Saccharomyces cerevisiae haploids and diploids";
 RL Biochem. Biophys. Res. Commun. 205:1079-1085(1994).
 CC [1]- FUNCTION: TRANSCRIPTION FACTOR INVOLVED IN THE FORMATION OF
 PSEUDOHYPHAE AND HYPHAE. IT IS LIKELY TO PLAY A ROLE IN THE
 DEVELOPMENTAL SWITCH BETWEEN YEAST AND MYCELIAL FORMS. MAY BE
 INVOLVED IN A SIGNAL TRANSDUCTION SYSTEM, STRENGTHENING THE
 POSSIBILITY OF A SEXUAL PHASE UP TO NOW UNDETECTED, AND SIMILAR TO
 THAT OF THE YEAST MATING PATHWAY.
 CC [2]- SUCELLULAR LOCATION: Nuclear.
 CC [3]- SIMILARITY: BELONGS TO THE STE12 TRANSCRIPTION FACTOR FAMILY.
 CC [4]- This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC DR EMBL, U15152; AAA64692.1; ALT_TERM.
 CC EMBL, U15152; AAA64692.1; ALT_TERM.
 CC PIR, A54767; A54767.
 DR InterPro: IPR003120; TF_STE.
 DR Pfam: PF02200; STE_1.
 DR SMART, SM00424; STE_1.
 KM Transcription regulation; DNA-binding; Activator; Nuclear protein.
 FT DNA BIND 44 154 BY SIMILARITY.
 FT DOMAIN 455 458 POLY-ALA.
 FT DOMAIN 555 563 POLY-GLN.
 FT DOMAIN 574 586 POLY-GLN.
 FT DOMAIN 593 596 POLY-LYS.
 FT CONFLICT 462 462 I -> M (IN REF. 2).
 FT CONFLICT 510 510 A -> T (IN REF. 2).
 FT CONFLICT 610 610 K -> N (IN REF. 2).
 SQ SEQUENCE 656 AA; 72978 MW; B55594BF019A9B09 CRC64;

DC12_ECOLI
 ID DC12_ECOLI STANDARD; PRT; 713 AA.
 AC P52055; P78299;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysine decarboxylase, constitutive (BC 4.1.1.18) (LDC).
 GN LDC OR LDC OR LDC OR B0186.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxId=562;
 RN [1]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=97369816; PubMed=9226257;
 RA Kikuchi Y., Kojima H., Tanaka T., Takatsuma Y., Kamio Y.;
 RT "Characterization of a second lysine decarboxylase isolated from
 Escherichia coli";
 RL J. Bacteriol. 179:4486-4492(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-40 AND 415-456.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=97480927; PubMed=9339543;
 RA Yamamoto Y., Miwa Y., Miyoshi K., Furuyama J., Ohmori H.;
 RT "The Escherichia coli ldc gene encodes another lysine decarboxylase,
 probably a constitutive enzyme";
 RL Genes Genet. Syst. 72:167-172(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:123-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RX Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of the
 4.0 - 6.0 min (189,987 - 281,416bp) region";
 RL Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung B.,
 RA Davis K., Federpiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
 RA Lahekari D., Lew H., Lin D., Namath A., Oelfer P., Roberts D.,
 RA Davis R.W.;
 RT Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 555-713 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=98389317; PubMed=9723924;
 RA Pichoff S., Alibaud L., Guedant A., Castanie M.-P., Bouche J.-P.;
 RT "An Escherichia coli gene (yaeO) suppresses temperature-sensitive
 mutations in essential genes by modulating Rho-dependent
 transcription termination";
 RL Mol. Microbiol. 29:859-869(1998).
 CC [1]- FUNCTION: LDC IS CONSTITUTIVELY BUT WEAKLY EXPRESSED UNDER VARIOUS
 CONDITIONS. OPTIMUM ACTIVITY IS ACHIEVED BETWEEN PH 6.2 TO 8.0.
 CC [2]- CATALYTIC ACTIVITY: L-lysine = cadaverine + CO(2).
 CC [3]- COFACTOR: Pyridoxal phosphate.
 CC [4]- SUBUNIT: HOMODECAMER.
 CC [5]- SIMILARITY: BELONGS TO FAMILY 1 OF ORNITHINE, LYSINE, AND ARGININE
 DECARBOXYLASES.
 CC [6]- This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; D87518; BAA21656.1; -
 DR EMBL; D49445; BAA08426.1; -
 DR EMBL; AE000128; AAC73297.1; -
 DR EMBL; D85316; BAA77861.1; -
 DR EMBL; U70214; AAB08615.1; -
 DR EMBL; Z50870; CAA90749.1; -
 DR PIR; B64743; B64743.
 DR HSSP; P43099; 1C4K.
 DR BcoGene; EG33219; 1dc.
 DR InterPro; IPR000310; Decarboxylase.
 DR InterPro; IPR005308; OKR_DC_1_N.
 DR Pfam; PF01276; OKR_DC_1; 1.
 DR Pfam; PF03711; OKR_DC_1_C; 1.
 DR Pfam; PF03709; OKR_DC_1_N; 1.
 DR PROSITE; PS00703; OKR_DC_1; 1.
 DR Lyase; Decarboxylase; Pyridoxal phosphate; Multigene family;
 KW Complete proteome.
 FT BINDING 367 PYRIDOXAL PHOSPHATE.
 FT CONFLICT 284 S -> T (IN REF. 2).
 FT CONFLICT 314 N -> F (IN REF. 2).
 FT CONFLICT 411 T -> S (IN REF. 2).
 FT CONFLICT 413 AA -> R (IN REF. 2).
 FT CONFLICT 498 M -> I (IN REF. 2).
 FT CONFLICT 673 L -> I (IN REF. 2).
 FT SEQUENCE 713 AA; 80590 MW; 45320609744ABDF CRC64;

Query Match 78.84; Score 26; DB 1; Length 713;
 Best Local Similarity 66.74; Pred. NO. 3.3e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 Db 118 RQYTD 123

RESULT 31
 ID ODP1_ECOLI STANDARD; PRT; 886 AA.
 AC P06958; P78049;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 PY Pyruvate dehydrogenase E1 component (EC 1.2.4.1).
 ACCE OR B0114 OR Z0124 OR E050118.
 OS *Escherichia coli*, and
 OS *Escherichia coli* O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OX NCBI_TaxID=562, 83334;
 (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=83209630; PubMed=6343085;
 RA Stephens P.B., Darlison M.G., Lewis H.M., Guest J.R.;
 RT "The pyruvate dehydrogenase complex of *Escherichia coli* K12.
 RT Nucleotide sequence encoding the pyruvate dehydrogenase component.";
 RL Eur. J. Biochem. 133:155-162(1983).
 (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=94261430; PubMed=8202364;
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RT "Systematic sequencing of the *Escherichia coli* genome: analysis of
 RT the 2.4.4.1 min (110,917-193,643 bp) region.";
 RL Nucleic Acids Res. 22:1637-1639(1994).
 (3)
 RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:12453-12474(1997).
 (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Pocamonsis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 (5)
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 (6)
 RP SEQUENCE OF 1-12.
 RC STRAIN=K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of *Escherichia coli* K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 (7)
 RN X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
 RP MEDLINE=21952337; PubMed=11955070;
 RA Arjunan P., Nemeria N., Brunschlag A., Chandrasekhar K., Sax M.,
 RA Yan Y., Jordan F., Guest J.R., Furey W.;
 RT "Structure of the pyruvate dehydrogenase multienzyme complex E1
 RT component from *Escherichia coli* at 1.85 Å resolution.";
 RL Biochemistry 41:5213-5221(2002).
 CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
 CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
 CC multiple copies of three enzymatic components: pyruvate
 CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
 CC lipoamide dehydrogenase (E3).
 CC -1- CATALYTIC ACTIVITY: Pyruvate + lipoamide = S-
 CC acetyldihydrolipoamide + CO(2).
 CC -1- COFACTOR: Thiamine pyrophosphate and magnesium ions.
 CC -1- SUBUNIT: Homodimer.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR PDB; 1L6A; 24-JUL-02.
 DR SWISS-2DPAGE; P06958; COLI.
 DR ECODBASE; F099.0; 6TH EDITION.
 DR Ecogene; EG10024; acef.
 DR InterPro; IPR004660; Acef.
 DR InterPro; IPR005474; Transketolase_N.
 DR Pfam; PF00456; transketolase; 1.
 DR TIGRFAMs; TIGR00759; acef; 1.
 DR Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;
 KM Magnesium; Complete proteome; 3D-structure.
 FT INIT_MET 0
 FT METAL 230
 FT METAL 260
 FT METAL 262
 FT METAL 262
 FT METAL 145
 FT CONFLICT 145
 FT CONFLICT 275
 SQ SEQUENCE 886 AA; 99537 MW; B8008F642184EC8 CRC64;

Query Match
 Best Local Similarity 78.8%; Score 26; DB 1; Length 886;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 543 QOYTDQ 548

RESULT 32
 YK03_YEAST
 ID YK03_YEAST STANDARD; PRT; 1038 AA.
 AC P36097;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Hypothetical 118.9 kDa protein in PTM-1XRI intergenic region.
 GN YK033W OR YK0246.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=6288C;
 RX MEDLINE=94262309; PubMed=8203146;
 RA Purrelle B., Skala J., van Dyck L., Goffeau A.;
 RT "Analysis of an 11.7 kb DNA fragment of chromosome XI reveals a new
 tRNA gene and four new open reading frames including a leucine zipper
 protein and a homologue to the yeast mitochondrial regulator ABP2.";
 RT Yeast 10:125-130 (1994).
 [2]
 RP SEQUENCE OF 913-1038 FROM N.A.
 RA Rieger M.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X71622; -; NOT ANNOTATED_CDS.
 DR EMBL; Z28033; CA81868.1; -.
 DR PIR; S37854; S37854.
 DR SCD; S0001516; YK033W.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 211
 FT TRANSMEM 237
 FT TRANSMEM 238
 FT TRANSMEM 594
 FT TRANSMEM 610
 FT TRANSMEM 869
 FT TRANSMEM 885
 SQ SEQUENCE 1038 AA; 118888 MW; D068AEC2F4556A72 CRC64;

Query Match
 Best Local Similarity 78.8%; Score 26; DB 1; Length 1038;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 833 KEYTDQ 838

RESULT 33
 Y16A_BP74
 ID Y16A_BP74 STANDARD; PRT; 50 AA.
 AC 09TOTO;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical 5.9 kDa protein in as1A-arn intergenic region.
 GN Y16A OR AS1A.1A.
 OS Bacteriophage T4.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses.
 OX NCBI_TaxID=10665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kutler E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
 RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
 RT "Bacteriophage T4 genome analysis.";
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF158101; AAD42540.1; -.
 KW Hypothetical protein.
 KM
 SQ SEQUENCE 50 AA; 5935 MW; C0A6966A5106A436 CRC64;

Query Match
 Best Local Similarity 75.8%; Score 25; DB 1; Length 50;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 30 ENYTDQ 35

RESULT 34
 Y122_CABEL
 ID Y122_CABEL STANDARD; PRT; 112 AA.
 AC P03935;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hypothetical 12.2 kDa protein from transposable element Tc1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=83246543; PubMed=6306578;
 RX Rosenzweig B., Liao L.W., Hirsch D.;
 RA "Sequence of the C. elegans transposable element Tc1.";
 RT Nucleic Acids Res. 11:4201-4209 (1983).
 RL
 CC -1- SIMILARITY: 34% IDENTITY TO HYPOTHETICAL PROTEIN FROM TRANSPOSABLE
 CC ELEMENT UHU1 OF DROSOPHILA HETERONEURA.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X01005; CAA25499.1; -
 DR PIR; A04521; OOKWTB.
 KM Hypothetical protein; Transposable element.
 SQ SEQUENCE 112 AA; 12249 MW; BCA0554A9DD9032B CRC64;

Query Match 75.8%; Score 25; DB 1; Length 112;
 Best Local Similarity 66.7%; Pred. No. 75;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 17 QOYTDQ 22

RESULT 35

ID PRO1_DICDI STANDARD; PRT; 126 AA.
 AC P26199;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Profilin I.
 GN PROA.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 OX NCBI_TaxId=44689;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX2;
 RX MEDLINE=92226170; PubMed=1725525;
 RA Haugwitz M., Noegel A.A., Rieger D., Lottepeich F., Schleicher M.;
 RT "Dictyostelium discoideum contains two profilin isoforms that differ
 RT in structure and function."
 RL J. Cell Sci. 100:481-489(1991).

CC -1- FUNCTION: Binds to actin and affects the structure of the
 CC cytoskeleton. At high concentrations, profilin prevents the
 CC polymerization of actin, whereas it enhances it at low
 CC concentrations. By binding to PIP2, it inhibits the formation of
 CC IP3 and DG.
 CC -1- SUBUNIT: OCCURS IN MANY KINDS OF CELLS AS A COMPLEX WITH MONOMERIC
 CC ACTIN IN A 1:1 RATIO.

-1- SIMILARITY: BELONGS TO THE PROFILIN FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X61581; CAA43781.1; -
 DR PIR; A53255; FAD01.
 DR HSSP; Q42449; 1AOK.
 DR DICTYDB; DD01035; PROA.
 DR InterPro; IPR002097; Profilin.
 DR InterPro; IPR005455; Profilin_plant.
 DR Pfam; PF00235; Profilin; 1.
 DR PRINTS; PR00392; PROFILIN.
 DR PRINTS; PR00392; PROFILIN.
 DR SMART; SM00392; PROF. 1.
 DR PROSITE; PS00414; PROFILIN; 1.
 KM Actin-binding, Cytoskeleton, Multigene family.
 SQ SEQUENCE 126 AA; 13064 MW; 0A92CAD75D0CE370 CRC64;

Query Match 75.8%; Score 25; DB 1; Length 126;

Best Local Similarity 66.7%; Pred. No. 85;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 4 QOYTDQ 9

RESULT 36

ID Y896_VIBVU STANDARD; PRT; 181 AA.
 AC Q8DDR1;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical UPF0241 protein V10896.
 GN V10896.
 OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrrio.
 OX NCBI_TaxId=672;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMCP6;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.B.;
 RT "Complete genome sequence of Vibrio vulnificus CMCP6";
 RT Submitted (DEC-2002) to the EMBL/Genbank/DBSJ databases.
 CC -1- SIMILARITY: Belongs to the UPF0241 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE016800; AA009399.1; -
 DR HAMAP; MF_01058; -; 1.
 DR InterPro; IPR007336; DUF414.
 DR Pfam; PF04220; DUF414; 1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 181 AA; 20568 MW; ICP629C2FPA74BF CRC64;

Query Match 75.8%; Score 25; DB 1; Length 181;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 124 QOYTDQ 129

RESULT 37

ID PDX2_HUMAN STANDARD; PRT; 198 AA.
 AC P32119; P31945; P35701; Q92763;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Peroxiredoxin 2 (EC 1.11.1.-) (Thioredoxin peroxidase 1) (Thioredoxin-
 DE dependent peroxide reductase 1) (Thiol-specific antioxidant protein)
 DE (TSA) (PRP) (Natural killer cell enhancing factor B) (NKFB-B).
 GN PDX2 OR TDPX1 OR NKFB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

DR NCBI_TaxId=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94193012; PubMed=8144038;

RA Lim Y.-S., Cha M.-K., Kim H.-K., Kim I.-H.,
 RT "The thiol-specific antioxidant protein from human brain: gene
 RT cloning and analysis of conserved cysteine regions.",
 RL Gene 140:279-284(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9429283; PubMed=8026862;
 RA Shaw H., Butterfield L.H., Chiu R., Kim A.,
 RT "Cloning and sequence analysis of candidate human natural killer-
 RT enhancing factor genes",
 RL Immunogenetics 40:129-134(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2238257; PubMed=12477932;
 RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shennen C.M., Schler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stanclevon M.J., Usdin T.B., Toshiyuki S., Carninci P., Scheetz T.E.,
 RA Brownstein M.J., Petros M.B., Bonaldo M.F., Casavant T.L., Prange C.,
 RA Rahn S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 35-198 FROM N.A.
 RX Oberbauer I.,
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 17-25, 140-150 AND 163-185.
 RX TISSUE=keratinocytes;
 RC MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., Van Damme J., Puype W., Gesser B., Celis J.E.,
 RA Vandekekerckhove J.,
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.",
 RL Electrophoresis 13:960-969(1992).
 RN [6]
 RP SEQUENCE OF 17-26 AND 93-103.
 RX TISSUE=erythrocyte;
 RC MEDLINE=94147970; PubMed=8313871;
 RA Golz O., Hughes G.U., Frutiger S., Paquet N., Bairoch A.,
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Appel R.D., Walzer C.,
 RA Balant L., Hochstrasser D.F.,
 RT "Plasma and red blood cell protein maps: update 1993.",
 RL Electrophoresis 14:1223-1231(1993).
 RN [7]
 RP SEQUENCE OF 17-26, 111-135 AND 140-157.
 RX TISSUE=Colon carcinoma;
 RC MEDLINE=97295306; PubMed=9150948;
 RA Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.,
 RT "A two-dimensional gel database of human colon carcinoma proteins.",
 RL Electrophoresis 18:605-613(1997).
 RN [8]
 RP OVEROXIDATION OF CYS-51.
 RX MEDLINE=22028064; PubMed=11904290;
 RA Rabilloud T., Heller M., Gaennier F., Luche S., Rey C., Aebersold R.,
 RA Benamed M., Lousiot P., Linaudi J.,
 RT "Proteomic analysis of cellular response to oxidative stress.
 RT Evidence for in vivo overoxidation of peroxiredoxins at their active
 RT site.",
 RL J. Biol. Chem. 277:19396-19401(2002).

CC -1- FUNCTION: Involved in redox regulation of the cell. Reduces
 CC peroxides with reducing equivalents provided through the
 CC thioredoxin system. It is not able to receive electrons from
 CC glutaredoxin. May play an important role in eliminating peroxides
 CC generated during metabolism. Might participate in the signaling
 CC cascades of growth factors and tumor necrosis factor-alpha by
 CC regulating the intracellular concentrations of H(2)O(2).
 CC -1- FUNCTION: Enhances natural killer (NK) cells activity.
 CC -1- SUBUNIT: Homodimer; disulfide-linked, upon oxidation (by
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- MISCELLANEOUS: The active site is the redox-active Cys-51 oxidized
 CC to Cys-SOH. Cys-SOH rapidly reacts with Cys-172-SH of the other
 CC subunit to form an intermolecular disulfide with a concomitant
 CC homodimer formation. The enzyme may be subsequently regenerated by
 CC reduction of the disulfide by thioredoxin.
 CC -1- MISCELLANEOUS: Irreversibly inactivated by overoxidation of Cys-51
 CC (to Cys-SO(2)H and Cys-SO(3)H) upon oxidative stress.
 CC -1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, Z22548; CA80269.1; -;
 CC EMBL, L19185; AA50465.1; -;
 CC EMBL, BC000452; AAH00452.1; -;
 CC EMBL, BC003022; AAH03022.1; -;
 CC EMBL, X82321; CA57764.1; -;
 CC PIR, I68897; I68897.
 CC PDB, 1QWV; 28-JUL-00.
 CC SWISS-2DPAGE, P32119; HUMAN.
 CC Aarhuus/Ghent-2DPAGE, 6116; IEF.
 CC Genew; HGNC:9353; PRDX2.
 CC MIM: 600538; -;
 CC DR GO, GO:0005488; F:electron transporter activity; TAS.
 CC DR GO, GO:0008379; F:thioredoxin peroxidase activity; TAS.
 CC DR InterPro, IPR000866; Ahpc-TSA.
 CC DR Pfam, PF00578; Ahpc-TSA; 1.
 CC KW Antioxidant; Peroxidase; Oxidoreductase; Redox-active center;
 CC 3D-structure.
 CC FT ACT_SITE 51 51 CYSTEINE SULFENIC ACID INTERMEDIATE
 CC (BY SIMILARITY).
 CC FT DISULFID 51 51 INTERCHAIN (WITH C-172) (IN
 CC LINKED FORM) (BY SIMILARITY).
 CC FT DISULFID 172 172 INTERCHAIN (WITH C-51) (IN
 CC LINKED FORM) (BY SIMILARITY).
 CC FT CONFLICT 59 66 SNRADER -> TYTKRTSA (IN REF. 1).
 CC FT CONFLICT 82 82 T -> N (IN REF. 2).
 CC FT CONFLICT 105 105 A -> G (IN REF. 2).
 CC FT CONFLICT 120 120 T -> N (IN REF. 1).
 CC FT CONFLICT 175 175 G -> A (IN REF. 1).
 CC FT CONFLICT 180 180 S -> R (IN REF. 1).
 CC SQ SEQUENCE 198 AA; 21892 MW; 1AC781D908B3B46 CAC64;
 CC
 CC Query Match 75.8%; Score 25; DB 1; Length 198;
 CC Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 CC Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC QY 2 QYTDQ 6
 CC |||||
 CC DB 163 QYTDH 167
 CC
 CC RESULT 38
 CC ID PDX2_MOUSE STANDARD; PRT; 198 AA.
 CC AC Q61171; Q60796;
 CC DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Peroxidoxin 2 (EC 1.11.1.-) (Thioredoxin peroxidase 1) (Thioredoxin-dependent peroxidase reductase 1) (Thiol-specific antioxidant protein) (TSA).
 GN PDX2 OR TDPX1 OR TPX.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA1B/C; TISSUE=Brain;
 RX MEDLINE=97238626; PubMed=9115640;
 RA Ichimura S., Davis J.G., O'Rourke D.M., Katsumata M., Greene M.I.,
 RT "Murine thioredoxin peroxidase delays neuronal apoptosis and is expressed in areas of the brain most susceptible to hypoxic and ischemic injury.";
 RT DNA Cell Biol. 16:311-321(1997).
 RT [2]
 SEQUENCE FROM N.A.
 RA STRAIN=129;
 RL Oberbauer I.;
 RP Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Chae H.Z., Kim H., Rhee S.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Involved in redox regulation of the cell. Reduces peroxidases with reducing equivalents provided through the thioredoxin system. It is not able to receive electrons from glutaredoxin. May play an important role in eliminating peroxides generated during metabolism. Might participate in the signaling cascades of growth factors and tumor necrosis factor-alpha by regulating the intracellular concentrations of H(2)O(2).
 CC -1- SUBUNIT: Homodimer; disulfide-linked, upon oxidation (By similarity).
 CC -1- SIMILARITY: Homodimer; disulfide-linked, upon oxidation (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in bone marrow. High levels also found in heart, brain, kidney and skeletal muscle. Lower levels in liver, lung and thymus.
 CC -1- MISCELLANEOUS: The active site is the redox-active Cys-51 oxidized to Cys-SOH. Cys-SOH rapidly reacts with Cys-172-SH of the other subunit to form an intermolecular disulfide with a concomitant homodimer formation. The enzyme may be subsequently regenerated by reduction of the disulfide by thioredoxin (By similarity).
 CC -1- MISCELLANEOUS: Irreversibly inactivated by overoxidation of Cys-51 (to Cys-SO2H and Cys-SO3H) upon oxidative stress (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE AHPc/TSA FAMILY.
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U51679; AAB01941.1; -
 CC EMBL: X82067; CAAS7566.1; -
 CC EMBL: U20611; AAA69475.1; -
 CC HSSP: P30041; 1PRX.
 CC DR SWISS-2DPAGE: O61171; MOUSE.
 CC MGD: MGI:109486; Pdx2.
 CC GO: GO:0008189; F:apoptosis inhibitor activity; IDA.
 CC GO: GO:0006916; P:anti-apoptosis; IDA.
 CC InterPro: IPR000866; AHPc-TSA.
 CC Pfam: PF00578; AHPc-TSA; 1.
 CC Antioxidant; Peroxidase; Oxidoreductase; Redox-active center.
 FT ACT_SITE 51 51
 FT CISTEINE SUPERNIC ACID INTERMEDIATE
 FT (BY SIMILARITY).
 FT DISULFID 51 51
 FT INTERCHAIN (WITH C-172) (IN

FT FT DISULFID 172 172 LINKED FORM (BY SIMILARITY).
 FT INTERCHAIN (WITH C-51) (IN
 FT CONFLICT 97 97 LINKED FORM (BY SIMILARITY).
 FT CONFLICT 182 182 G -> A (IN REF. 3).
 FT SEQUENCE 198 AA; 21778 MW; FE216F5426F7174D CRC64;
 SQ
 Query Match 75.8%; Score 25; DB 1; Length 198;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 QYTDQ 6
 Db 163 QYTDQ 167
 RESULT 39
 PDX2 RAT STANDARD; PRT; 198 AA.
 ID PDX2 RAT
 AC P35704;
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DE 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Peroxidoxin 2 (EC 1.11.1.-) (Thioredoxin peroxidase 1) (Thioredoxin-dependent peroxidase reductase 1) (Thiol-specific antioxidant protein) (TSA).
 GN PDX2 OR TDPX1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID:10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=9431629; PubMed=80411738;
 RA Chae H.Z., Robinson K., Poole L.B., Church G., Storz G., Rhee S.G.;
 RL "Cloning and sequencing of thiol-specific antioxidant from mammalian brain: alyl hydroperoxide reductase and thiol-specific antioxidant define a large family of antioxidant enzymes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:7017-7021(1994).
 CC -1- FUNCTION: Involved in redox regulation of the cell. Reduces peroxidases with reducing equivalents provided through the thioredoxin system. It is not able to receive electrons from glutaredoxin. May play an important role in eliminating peroxides generated during metabolism. Might participate in the signaling cascades of growth factors and tumor necrosis factor-alpha by regulating the intracellular concentrations of H(2)O(2).
 CC -1- SUBUNIT: Homodimer; disulfide-linked, upon oxidation (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- MISCELLANEOUS: The active site is the redox-active Cys-51 oxidized to Cys-SOH. Cys-SOH rapidly reacts with Cys-172-SH of the other subunit to form an intermolecular disulfide with a concomitant homodimer formation. The enzyme may be subsequently regenerated by reduction of the disulfide by thioredoxin (By similarity).
 CC -1- MISCELLANEOUS: Irreversibly inactivated by overoxidation of Cys-51 (to Cys-SO2H and Cys-SO3H) upon oxidative stress (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE AHPc/TSA FAMILY.
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U06099; AAA19959.1; -
 CC PIR: A57716; A57716.
 CC HSSP: P30041; 1PRX.
 CC InterPro: IPR000866; AHPc-TSA.
 CC Pfam: PF00578; AHPc-TSA; 1.
 CC Antioxidant; Peroxidase; Oxidoreductase; Redox-active center.

FT ACT_SITE 51 51 CYSTEINE SULFENIC ACID INTERMEDIATE
 FT DISULFID 51 51 (BY SIMILARITY).
 FT INTERCHAIN (WITH C-172) (IN
 FT DISULFID 172 172 LINKED FORM) (BY SIMILARITY).
 FT INTERCHAIN (WITH C-51) (IN
 FT SEQUENCE 198 AA; 21784 MW; FCBAD90E9C447B CRC64;
 LINKED FORM) (BY SIMILARITY).

Query Match
 Best Local Similarity 75.8%; Score 25; DB 1; Length 198;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QYTDQ 6
 Db 163 QYTDE 167

RESULT 40

PDX2_BOVIN STANDARD; PRT; 199 AA.

Q9BG13;

28-FEB-2003 (Rel. 41, Created)

28-FEB-2003 (Rel. 41, Last sequence update)

15-SEP-2003 (Rel. 42, Last annotation update)

Peroxiredoxin 2 (BC 1.11.1.-).

PRDX2.

Bos taurus (Bovine).

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

NCBI_TaxID=9913;

[1]

SEQUENCE FROM N.A.

TISSUE=Liver;

Leyens G., Donay I., Knoops B.;

"Cloning of 4 new bovine peroxiredoxins, and screening of the complete
 peroxiredoxin family in different bovine tissues."

Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: Involved in redox regulation of the cell. Reduces
 peroxides with reducing equivalents provided through the
 thioredoxin system. It is not able to receive electrons from
 glutaredoxin. May play an important role in eliminating peroxides
 generated during metabolism. Might participate in the signaling
 cascades of growth factors and tumor necrosis factor-alpha by
 regulating the intracellular concentrations of H(2)O(2).

-1- SUBUNIT: Homodimer; disulfide-linked, upon oxidation (By
 similarity).

-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

-1- MISCELLANEOUS: The active site is the redox-active Cys-52 oxidized
 to Cys-SOH. Cys-SOH rapidly reacts with Cys-173-SH of the other
 subunit to form an intermolecular disulfide with a concomitant
 homodimer formation. The enzyme may be subsequently regenerated by
 reduction of the disulfide by thioredoxin (By similarity).

-1- MISCELLANEOUS: Irreversibly inactivated by overoxidation of Cys-52
 (to Cys-SO(2)H and Cys-SO(3)H) upon oxidative stress (By similarity).

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

-1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.

FT DISULFID 173 173 LINKED FORM) (BY SIMILARITY).
 FT INTERCHAIN (WITH C-52) (IN
 FT SEQUENCE 199 AA; 21946 MW; 5F256CE54090E2DE CRC64;
 LINKED FORM) (BY SIMILARITY).

Query Match
 Best Local Similarity 75.8%; Score 25; DB 1; Length 199;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QYTDQ 6
 Db 164 QYTDE 168

Search completed: August 20, 2003, 09:30:46
 Job time : 6.46392 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:19:04 ; Search time 17.4433 Seconds
(without alignments)
88.763 Million cell updates/sec

Title: US-09-991-433-7

Perfect score: 33

Sequence: 1 QOYTDQ 6

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

1 number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	138	12	Q85155 human parvo
2	33	100.0	141	12	Q85171 human parvo
3	33	100.0	141	12	Q85166 human parvo
4	33	100.0	141	12	Q85146 human parvo
5	33	100.0	142	12	Q85168 human parvo
6	33	100.0	142	12	Q85173 human parvo
7	33	100.0	144	12	Q85138 human parvo
8	33	100.0	145	12	Q85181 human parvo
9	33	100.0	145	12	Q85161 human parvo
10	33	100.0	146	12	Q85185 human parvo
11	33	100.0	146	12	Q85158 human parvo
12	33	100.0	147	12	Q85142 human parvo
13	33	100.0	148	12	Q85150 human parvo
14	33	100.0	151	12	Q85131 human parvo
15	33	100.0	151	12	Q85177 human parvo
16	33	100.0	153	12	Q85123 human parvo

17	33	100.0	157	12	Q85196 human parvo
18	33	100.0	162	12	Q85135 human parvo
19	33	100.0	546	12	Q913X0 human parvo
20	33	100.0	546	12	Q913W7 human parvo
21	33	100.0	554	12	Q9PZS9 human parvo
22	33	100.0	554	12	Q90201 human parvo
23	33	100.0	554	12	Q65790 human parvo
24	33	100.0	554	12	Q912B7 human eryth
25	33	100.0	554	12	Q9JGP7 human eryth
26	33	100.0	554	12	Q9WKJ9 human parvo
27	33	100.0	554	12	Q8UN54 human parvo
28	33	100.0	760	12	Q9PZT8 human parvo
29	33	100.0	761	12	Q9PZU0 human parvo
30	33	100.0	765	12	Q9PZT6 human parvo
31	33	100.0	769	12	Q9PZT4 human parvo
32	33	100.0	773	12	Q913X1 human parvo
33	33	100.0	773	12	Q913W8 human parvo
34	33	100.0	781	12	Q8UN56 human parvo
35	33	100.0	781	12	Q65789 human parvo
36	33	100.0	781	12	P90223 human parvo
37	33	100.0	781	12	Q85191 human parvo
38	33	100.0	781	12	P90221 human parvo
39	33	100.0	781	12	P89318 human parvo
40	33	100.0	781	12	Q8UYE3 human parvo
41	33	100.0	781	12	P90224 human parvo
42	33	100.0	781	12	Q9PZT0 human parvo
43	33	100.0	781	12	P89316 human parvo
44	33	100.0	781	12	P89320 human parvo
45	33	100.0	781	12	Q85117 human parvo

ALIGNMENTS

RESULT 1
ID Q85155 PRELIMINARY; PRT; 138 AA.
AC Q85155;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Viral protein 1 (fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_Taxid=10798;
RN (1)
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Giegler A., Caesinotti P., Siegl G., Wolf H., Modrow S.;
RT "xxxSequence variability among different parvovirus B19 isolates.";
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70564; CAA94477.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA, 15237 MW, 5899FB8879A3B68D CRC64;
Query Match 100.0%; Score 33; DB 12; Length 138;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYTDQ 6
DB 111 QOYTDQ 116
RESULT 2
ID Q85171 PRELIMINARY; PRT; 141 AA.
AC Q85171;
DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 270580; CAA94493.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 141
FT NON_TER 141 141
SQ SEQUENCE 141 AA; 15770 MW; C9E92572A78C6C27 CRC64;
Query Match 100.0%; Score 33; DB 12; Length 141;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 113 QOYTDQ 118
RESULT 3
ID 085166 PRELIMINARY; PRT; 141 AA.
AC 085166;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 270575; CAA94488.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 141
FT NON_TER 141 141
SQ SEQUENCE 141 AA; 15785 MW; 4115D3D915751757 CRC64;
Query Match 100.0%; Score 33; DB 12; Length 141;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 113 QOYTDQ 118
RESULT 4
ID 085146 PRELIMINARY; PRT; 141 AA.
AC 085146;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 270555; CAA94467.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 141
FT NON_TER 141 141
SQ SEQUENCE 141 AA; 15712 MW; 32F6B23E7E6B8583 CRC64;
Query Match 100.0%; Score 33; DB 12; Length 141;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 110 QOYTDQ 115
RESULT 5
ID 085168 PRELIMINARY; PRT; 142 AA.
AC 085168;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 270577; CAA94490.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 142
FT NON_TER 142 142
SQ SEQUENCE 142 AA; 15821 MW; 2B6E4D9A6784F8C4 CRC64;
Query Match 100.0%; Score 33; DB 12; Length 142;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 109 QOYTDQ 114
RESULT 6
ID 085173 PRELIMINARY; PRT; 142 AA.
AC 085173;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,

RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70582; CAA94495.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 142 142
 SQ SEQUENCE 142 AA; 15820 MW; 3A9E5E07C0BC0434 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 33; DB 12; Length 142;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YTDQ 6
 Db 109 00YTDQ 114

RESULT 7

AC 085138 PRELIMINARY; PRT; 144 AA.
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblutzki A., Gigler A., Casinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70547; CAA94459.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 144 144
 SQ SEQUENCE 144 AA; 16022 MW; 95D6EF37BA5997AE CRC64;

Query Match
 Best Local Similarity 100.0%; Score 33; DB 12; Length 144;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 00YTDQ 6
 111 00YTDQ 116

RESULT 8

ID 085181 PRELIMINARY; PRT; 145 AA.
 AC 085181;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblutzki A., Gigler A., Casinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70550; CAA94503.1; -
 FT NON_TER 1 146 146
 SQ SEQUENCE 146 AA; 16283 MW; 77DCA9E25005FF8E CRC64;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 1
 RT NON_TER 145 145
 SQ SEQUENCE 145 AA; 16210 MW; 6B45A1B1B9E923C4 CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 33; DB 12; Length 145;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YTDQ 6
 Db 113 00YTDQ 118

RESULT 9

ID 085161 PRELIMINARY; PRT; 145 AA.
 AC 085161;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblutzki A., Gigler A., Casinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70570; CAA94483.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 145 145
 SQ SEQUENCE 145 AA; 16165 MW; B4FE95AB6E4D9A67 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 33; DB 12; Length 145;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YTDQ 6
 Db 109 00YTDQ 114

RESULT 10

ID 085185 PRELIMINARY; PRT; 146 AA.
 AC 085185;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblutzki A., Gigler A., Casinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70594; CAA94507.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 146 146
 SQ SEQUENCE 146 AA; 16283 MW; 77DCA9E25005FF8E CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 33; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|||||
DB 113 QOYTDQ 118

RESULT 11

ID 085158 PRELIMINARY; PRT; 146 AA.

AC 085158;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VPI.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1]

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giggler A., Cassinotti P., Siegl G.,

Wolff H., Modrow S.;

RT "XXSsequence variability among different parvovirus B19 isolates.";

RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z70557; CA94480.1; -

DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 146 146

SO SEQUENCE 146 AA; 16252 MW; 2B74FE95AB6E4D9A CRC64;

Query Match 100.0%; Score 33; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|||||
DB 109 QOYTDQ 114

RESULT 12

ID 085142 PRELIMINARY; PRT; 147 AA.

AC 085142;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VPI.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1]

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giggler A., Cassinotti P., Siegl G.,

Wolff H., Modrow S.;

RT "XXSsequence variability among different parvovirus B19 isolates.";

RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z70551; CA94463.1; -

DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 147 147

SO SEQUENCE 147 AA; 16424 MW; E4ECF2459B308BA9 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 147;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|||||

DB 112 QOYTDQ 117

RESULT 13

ID 085150 PRELIMINARY; PRT; 148 AA.

AC 085150;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VPI.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1]

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giggler A., Cassinotti P., Siegl G.,

Wolff H., Modrow S.;

RT "XXSsequence variability among different parvovirus B19 isolates.";

RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z70559; CA94471.1; -

DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 148 148

SO SEQUENCE 148 AA; 16539 MW; 8B14ECF2459B308B CRC64;

Query Match 100.0%; Score 33; DB 12; Length 148;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|||||
DB 112 QOYTDQ 117

RESULT 14

ID 085131 PRELIMINARY; PRT; 151 AA.

AC 085131;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VPI.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1]

SEQUENCE FROM N.A.

RX MEDLINE=96332516; PubMed=8760426;

RA Hemauer A., Von Poblotszki A., Giggler A., Cassinotti P., Siegl G.,

Wolff H., Modrow S.;

RT "Sequence variability among different parvovirus B19 isolates.";

RL J. Gen. Virol. 77:1781-1785 (1996).

DR EMBL: Z70540; CA94452.1; -

DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 151 151

SO SEQUENCE 151 AA; 16902 MW; 3CA74914B8E73A3E CRC64;

Query Match 100.0%; Score 33; DB 12; Length 151;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
|||||
DB 113 QOYTDQ 118

RESULT 15

085177 PRELIMINARY; PRT; 151 AA.
 ID 085177;
 AC 085177;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DE Viral protein 1 (Fragment).
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Gigler A., Caesinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270586; CAA94499.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 151
 FT NON_TER 151 151
 SQ SEQUENCE 151 AA; 16888 MW; 06C6A614A616DDA3 CRC64;
 Query Match 100.0%; Score 33; DB 12; Length 151;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 DB 113 QOYTDQ 118
 RESULT 16
 ID 085123 PRELIMINARY; PRT; 153 AA.
 AC 085123;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Gigler A., Caesinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "Sequence variability among different parvovirus B19 isolates."
 RL J. Gen. Virol. 77:11781-11785(1996).
 DR EMBL; 270532; CAA94444.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 153
 FT NON_TER 153 153
 SQ SEQUENCE 153 AA; 17131 MW; ECCA1F44020814EC CRC64;
 Query Match 100.0%; Score 33; DB 12; Length 153;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 DB 112 QOYTDQ 117
 RESULT 17
 ID 085196 PRELIMINARY; PRT; 157 AA.
 AC 085196;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)

DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Gigler A., Caesinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270603; CAA94518.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 157
 FT NON_TER 157 157
 SQ SEQUENCE 157 AA; 17591 MW; 1D42191887FCE03 CRC64;
 Query Match 100.0%; Score 33; DB 12; Length 157;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 DB 114 QOYTDQ 119
 RESULT 18
 ID 085135 PRELIMINARY; PRT; 162 AA.
 AC 085135;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Gigler A., Caesinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270544; CAA94456.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 162
 FT NON_TER 162 162
 SQ SEQUENCE 162 AA; 18005 MW; 3F51443566660F2 CRC64;
 Query Match 100.0%; Score 33; DB 12; Length 162;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYTDQ 6
 DB 113 QOYTDQ 118
 RESULT 19
 ID 0913X0 PRELIMINARY; PRT; 546 AA.
 AC 0913X0;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
 DE Major virus capsid protein VP2 (Fragment).
 GN VP2.
 OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RC SEQUENCE FROM N.A.
RA Hokynar K.; Soderlund-Venermo M., Ranki A., Kiviluoto O., Partio E.K., Hedman K.;
RT "A new parvovirus B19 genotype persistent in skin."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY044266; AAK5572.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 546
SQ SEQUENCE 546 AA; 59934 MW; 467BE468A67282E5 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOXTDQ 6
|||||
399 QOXTDQ 404

RESULT 20
ID Q913W7 PRELIMINARY; PRT; 546 AA.
AC Q913W7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Major virus capsid protein VP2 (Fragment).
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RC SEQUENCE FROM N.A.
RA Hokynar K.; Soderlund-Venermo M., Ranki A., Kiviluoto O., Partio E.K., Hedman K.;
RT "A new parvovirus B19 genotype persistent in skin."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY044266; AAK5575.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 546
SQ SEQUENCE 546 AA; 59910 MW; BBFLDD23695E3FCD CRC64;

Query Match 100.0%; Score 33; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOXTDQ 6
|||||
399 QOXTDQ 404

RESULT 21
ID Q9P2S9 PRELIMINARY; PRT; 554 AA.
AC Q9P2S9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VP2 capsid protein.
GN VP.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=HV;
FT NON_TER 554
SQ SEQUENCE 554 AA; 60853 MW; 98FE598FF20CB66F CRC64;

RA Gallinella G., Venturoli S.;
RT "B19 Genome Sequence and Structure Analysis."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF162273; AAD4615.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 554
SQ SEQUENCE 554 AA; 60833 MW; 7901FBD65A697E6 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOXTDQ 6
|||||
399 QOXTDQ 404

RESULT 22
ID Q90201 PRELIMINARY; PRT; 554 AA.
AC Q90201;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VP2 structural protein (Capsid protein VP2) (Fragment).
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RC SEQUENCE FROM N.A.
RA MEDLINE=97081188; PubMed=8922470;
RX Erdman D.D., Durigon E.L., Wang O.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the VP1/VP2 gene from multiple isolates."
RL J. Gen. Virol. 77:2767-2774(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MI, and N8;
RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K., Sasaki T.;
RT "Sequence of human parvovirus B19 isolates from patients with rheumatoid arthritis."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=N8;
RX MEDLINE=90218047; PubMed=2157807;
RA Umeke K., Nunoue T.;
RT "The genome type of human parvovirus B19 strains isolated in Japan during 1981 differs from types detected in 1986 to 1987: a correlation between genome type and prevalence."
RL J. Gen. Virol. 71:983-986(1990).
RN [4]
DR EMBL: U53595; AAB47453.1; -
DR EMBL: U53596; AAB47455.1; -
DR EMBL: U53593; AAB47449.1; -
DR EMBL: U53594; AAB47451.1; -
DR EMBL: U53587; AAB47457.1; -
DR EMBL: U53600; AAB47463.1; -
DR EMBL: U53601; AAB47465.1; -
DR EMBL: U53598; AAB47459.1; -
DR EMBL: U53599; AAB47461.1; -
DR EMBL: AB030693; BAA90290.1; -
DR EMBL: AB030673; BAA90268.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 554
SQ SEQUENCE 554 AA; 60853 MW; 98FE598FF20CB66F CRC64;

Query Match 100.0%; Score 33; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 399 QOYTDQ 404

RESULT 23

ID Q65790 PRELIMINARY; PRT; 554 AA.
AC Q65790;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VEN1;
Echovaria Mayo J.E., Erdman D.D.;
Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
EMBL: U1358; AAA83559.1; -
DR InterPro: IPR001403; Parvo_coat.
SQ SEQUENCE 554 AA; 60798 MW; BA89F2B293BE4E24 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 399 QOYTDQ 404

RESULT 24

ID Q912B7 PRELIMINARY; PRT; 554 AA.
AC Q912B7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE VP2 protein.
GN VP2.
OS Human erythrovirus V9.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=72197;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=V9;
Nguyen O.T.;
"Molecular cloning and sequencing of a novel human erythrovirus
genome: new species beside B19 in the genus Erythrovirus."
Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ249437; CAC80622.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat.
SQ SEQUENCE 554 AA; 60885 MW; BA7338CDAA80F336 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 399 QOYTDQ 404

RESULT 25

ID Q9UGP7 PRELIMINARY; PRT; 554 AA.
AC Q9UGP7;
DT 01-OCT-2000 (TRENBLrel. 15, Created)

DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Capsid protein VP2.
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rm;
RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K.,
Sasaki T.;
"Sequence of human parvovirus B19 isolates from patients with
rheumatoid arthritis."
Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB030694; BAA90293.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat.
SQ SEQUENCE 554 AA; 60839 MW; ODB958B33C73BE64 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 399 QOYTDQ 404

RESULT 26

ID Q9WKL9 PRELIMINARY; PRT; 554 AA.
AC Q9WKL9;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Structural protein VP2.
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20014169; PubMed=10548133;
RA Hemmer A., Beckenlehner K., Wolf H., Lang B., Modrow S.;
"Acute parvovirus B19 infection in connection with a flare of systemic
lupus erythematosus in a female patient."
J. Clin. Virol. 14:73-77 (1999).
DR EMBL: AF13323; AAC99439.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat.
SQ SEQUENCE 554 AA; 60913 MW; 98FB588FF205C66F CRC64;

Query Match 100.0%; Score 33; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
DB 399 QOYTDQ 404

RESULT 27

ID Q8JNS4 PRELIMINARY; PRT; 554 AA.
AC Q8JNS4;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;

RN [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN=D91.1;
 RX MEDLINE=22174902; PubMed=12186896;
 RA Servant A., Lapeche S., Lallemand F., Marinho V., De Saint Maur G.,
 RT "Genetic Diversity within Human Erythroviruses: Identification of
 RT Three Genotypes."
 RL J. Virol. 76:9124-9134 (2002).
 DR EMBL: AY083234; AAL91014.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60830 MW; A298575C083F24C8 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 626 QOYTDQ 404

RESULT 28
 ID Q9PZT8 PRELIMINARY; PRT; 760 AA.
 AC Q9PZT8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE VP1/2 (Fragment).
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RX NCBI_TaxID=10798;
 RN [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN=Kati1;
 RX MEDLINE=20191963; PubMed=10725428;
 RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
 RT "Integrity and full coding sequence of B19 virus DNA persisting in
 RT human synovial tissue."
 RL J. Gen. Virol. 81:1017-1025 (2000).
 DR EMBL: AF161224; AAD45912.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON TER 760 760
 SQ SEQUENCE 760 AA; 83403 MW; EA6EB0145E3A0B5A CRC64;

Query Match 100.0%; Score 33; DB 12; Length 760;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 626 QOYTDQ 631

RESULT 29
 ID Q9PZU0 PRELIMINARY; PRT; 761 AA.
 AC Q9PZU0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE VP1/2 (Fragment).
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RX NCBI_TaxID=10798;
 RN [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN=Kati1;
 RX MEDLINE=20191963; PubMed=10725428;

RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
 RA Partio E.K., Kontinen Y., Hedman K.;
 RT "Integrity and full coding sequence of B19 virus DNA persisting in
 RT human synovial tissue."
 RL J. Gen. Virol. 81:1017-1025 (2000).
 DR EMBL: AF161223; AAD45910.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON TER 761 761
 SQ SEQUENCE 761 AA; 83574 MW; 73A84C930D473530 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 761;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 626 QOYTDQ 631

RESULT 30
 ID Q9PZT6 PRELIMINARY; PRT; 765 AA.
 AC Q9PZT6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE VP1/2 (Fragment).
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RX NCBI_TaxID=10798;
 RN [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN=Kati3;
 RX MEDLINE=20191963; PubMed=10725428;
 RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
 RT "Integrity and full coding sequence of B19 virus DNA persisting in
 RT human synovial tissue."
 RL J. Gen. Virol. 81:1017-1025 (2000).
 DR EMBL: AF161225; AAD45915.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON TER 765 765
 SQ SEQUENCE 765 AA; 83998 MW; 89E254606DCDB8 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 765;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 DB 626 QOYTDQ 631

RESULT 31
 ID Q9PZT4 PRELIMINARY; PRT; 769 AA.
 AC Q9PZT4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE VP1/2 (Fragment).
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RX NCBI_TaxID=10798;
 RN [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN=Kati4;
 RX MEDLINE=20191963; PubMed=10725428;
 RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
 RA Partio E.K., Kontinen Y., Hedman K.;
 RT "Integrity and full coding sequence of B19 virus DNA persisting in

RT human synovial tissue."
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161226; AAD45917.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 769 769
SQ SEQUENCE 769 AA; 84578 MW; 0749D46E5CA7BB68 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 769;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
Db 626 QOYTDQ 631

RESULT 32
Q913X1

PRELIMINARY; PRT; 773 AA.

Q913X1; 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Minor virus capsid protein VPI (Fragment).

GN VPI.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=Lali;

RA Hokynar K.; Soderlund-Venermo M., Ranki A., Kivilluoto O., Partio E.K.,

RT "A new parvovirus B19 genotype persistent in skin."

RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL; AY044266; AAK95571.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 773 773

SQ SEQUENCE 773 AA; 85164 MW; D60CCA6F90B05378 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
Db 626 QOYTDQ 631

RESULT 33

Q913W8 PRELIMINARY; PRT; 773 AA.

Q913W8; 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Minor virus capsid protein VPI (Fragment).

GN VPI.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=HAM;

RA Hokynar K.; Soderlund-Venermo M., Ranki A., Kivilluoto O., Partio E.K.,

RT "A new parvovirus B19 genotype persistent in skin."

RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL; AY044268; AAK95574.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 773 773
SQ SEQUENCE 773 AA; 85140 MW; 2B86F3245F9CBE50 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
Db 626 QOYTDQ 631

RESULT 34

O8JNS6 PRELIMINARY; PRT; 781 AA.

O8JNS6; 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE VPI.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=D91.1;

RX MEDLINE=22174902; PubMed=12186896;

RA Serrant A., Lapeche S., Lallemand F., Marinho V., De Saint Maur G.,

RA Meritet J.F., Garbarg-Chenon A.;

RT "Genetic Diversity within Human Erythroviruses: Identification of

RT Three Genotypes."

RL J. Virol. 76:9124-9134(2002).

DR EMBL; AY088324; AAL91013.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

SQ SEQUENCE 781 AA; 86142 MW; 4EB71FEDD41FC8F3 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
Db 626 QOYTDQ 631

RESULT 35

O65789 PRELIMINARY; PRT; 781 AA.

O65789; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE VPI.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=VENI;

RA Echevarria Mayo J.E., Erdman D.D.;

RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.

DR EMBL; U13358; AAK83558.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

SQ SEQUENCE 781 AA; 85985 MW; E1D5FDE230935DD CRC64;

Query Match 100.0%; Score 33; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYTDQ 6
Db 626 QOYTDQ 631

Db 626 QOYTDQ 631

RESULT 36

P90223 PRELIMINARY; PRT; 781 AA.
 ID 085191
 AC P90223;
 DT 01-MAY-1997 (TEMBLrel. 03, Created)
 DT 01-MAY-1997 (TEMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USA2;
 RX MEDLINE=97081188; PubMed=8922470;
 RA Erdman D.D., Durigson E.L., Wang Q.Y., Anderson L.J.;
 "Genetic diversity of human parvovirus B19: sequence analysis of the
 VP1/VP2 gene from multiple isolates."
 J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38514; AAB47796.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86070 MW; C1D911E39CF04A88 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 781;

Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 Db 626 QOYTDQ 631

RESULT 37

085191 PRELIMINARY; PRT; 781 AA.
 ID 085191;
 AC 085191;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE Orf2 protein.
 OS ORF2.
 GN Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Hemmer A., Von Pöhlitz A., Giegler A., Cassinotti P., Siegl G.,
 Wolf H., Modrow S.;
 "XXS-sequence variability among different parvovirus B19 isolates."
 Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 RL EMBL; Z70599; CA94513.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86026 MW; AA02577B63BDB2D CRC64;

Query Match 100.0%; Score 33; DB 12; Length 781;

Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 Db 626 QOYTDQ 631

RESULT 38
 ID P90221 PRELIMINARY; PRT; 781 AA.
 AC P90221;
 DT 01-MAY-1997 (TEMBLrel. 03, Created)

DT 01-MAY-1997 (TEMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KOR2;
 RX MEDLINE=97081188; PubMed=8922470;
 RA Erdman D.D., Durigson E.L., Wang Q.Y., Anderson L.J.;
 "Genetic diversity of human parvovirus B19: sequence analysis of the
 VP1/VP2 gene from multiple isolates."
 J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38511; AAB47793.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86070 MW; 9FA830083F6F1357 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 781;

Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 Db 626 QOYTDQ 631

RESULT 39

P89318 PRELIMINARY; PRT; 781 AA.
 ID P89318;
 AC P89318;
 DT 01-MAY-1997 (TEMBLrel. 03, Created)
 DT 01-MAY-1997 (TEMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAP1;
 RX MEDLINE=97081188; PubMed=8922470;
 RA Erdman D.D., Durigson E.L., Wang Q.Y., Anderson L.J.;
 "Genetic diversity of human parvovirus B19: sequence analysis of the
 VP1/VP2 gene from multiple isolates."
 J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38509; AAB47791.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86064 MW; 3AE65CE6906339 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 781;

Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYTDQ 6
 Db 626 QOYTDQ 631

RESULT 40

08UYE3 PRELIMINARY; PRT; 781 AA.
 ID 08UYE3;
 AC 08UYE3;
 DT 01-OCT-2002 (TEMBLrel. 22, Created)
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE 7.5 kDa protein.
 OS Erythrovirus A6.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=182494;
 RN [1]

RP SEQUENCE FROM N.A.
RA Nguyen Q.T., Wong S., Brown K.E.;
RT "Identification and characterization of a second novel human
RT erythrovirus variant, A6."
RU Virology 0:0-0(2002).
DR EMBL; A1064475; AAL55418.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86277 MW; E73AB663E451128 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 781;

Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QOYTDQ 6

Db 626 QOYTDQ 631

Search completed: August 20, 2003, 09:24:00
Time : 18.443 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:19:04 ; Search time 13.6082 Seconds
(without alignments)
46.656 Million cell updates/sec

Title: US-09-991-433-8

Perfect score: 22

Sequence: 1 QOYQ 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq.19Jun03.*
1: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
2: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
3: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
4: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
5: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
6: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
7: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
8: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
9: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
10: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
11: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
12: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
13: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
14: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
15: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
16: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
17: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
18: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
19: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
20: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
21: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
22: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
23: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA2002.DAT.*
24: /SIDSI/gcgcdata/geneSeq/geneSeq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	10	AAU25791	Breast cancer-asso
2	22	100.0	13	AAW10722	Catalytic light ch
3	22	100.0	15	AAV25184	Foot and mouth dis
4	22	100.0	15	AAV70487	Synthesised peptid
5	22	100.0	15	AAV70488	Synthesised peptid
6	22	100.0	17	ABG73642	PERV antigenic pep
7	22	100.0	21	ABG75795	Bovine histatin 11
8	22	100.0	32	AAO07536	Human polypeptide
9	22	100.0	32	AAE32369	Human HBNF peptide

10	22	100.0	42	21	AAG47591	Arabidopsis thalia
11	22	100.0	45	21	AAG47590	Arabidopsis thalia
12	22	100.0	46	20	AAV26934	Human herpesvirus
13	22	100.0	46	22	AAV73930	Human herpesvirus
14	22	100.0	49	21	AAV47589	Arabidopsis thalia
15	22	100.0	51	23	ABU67216	G-protein coupled
16	22	100.0	51	24	ABP53997	Human muscarinic a
17	22	100.0	55	22	AAW84018	Human immune/haema
18	22	100.0	58	21	AAV12599	Zea mays protein f
19	22	100.0	58	21	AAV41160	Zea mays protein f
20	22	100.0	58	24	ABG75791	Bovine histatin 11
21	22	100.0	58	24	ABG75792	Bovine histatin 11
22	22	100.0	58	24	ABG75794	Bovine histatin 11
23	22	100.0	60	21	AAV15832	Arabidopsis thalia
24	22	100.0	64	21	ABG52856	Human liver peptid
25	22	100.0	64	22	ABV38011	Peptide #5517 enco
26	22	100.0	64	22	AAV56368	Human bone marrow
27	22	100.0	64	22	AAV71144	Human brain expres
28	22	100.0	64	22	AAV31425	Human peptide enco
29	22	100.0	64	23	ABG40941	Human peptide enco
30	22	100.0	66	23	ABP64187	Human ORF557, Hom
31	22	100.0	68	22	ABP35150	Human immune/haema
32	22	100.0	69	23	ABP31839	Human immune/haema
33	22	100.0	76	22	ABV69881	Human nucleas-11k
34	22	100.0	76	22	AAV56547	Drosophila melanog
35	22	100.0	78	16	AAV81957	Protonibacterium
36	22	100.0	78	22	AAO09512	ORF encoded by exo
37	22	100.0	78	23	ABP35150	Human polypeptide
38	22	100.0	78	24	ABR40617	Human ORF4123 prot
39	22	100.0	80	18	AAW27130	Zea mays oil trait
40	22	100.0	80	21	AAV70584	Equine rhinovirus
41	22	100.0	85	22	AAU17941	Salmonella Pathoge
42	22	100.0	86	21	AAV27732	Novel human respir
43	22	100.0	91	20	AAV06312	Arabidopsis thalia
44	22	100.0	93	21	AAV32702	Human interferon-e
45	22	100.0	96	19	AAV73526	Eucalyptus grandis

ALIGNMENTS

RESULT 1	AAU25791	standard; Peptide; 10 AA.
ID	AAU25791	
AC	AAU25791;	
XX		
DT	17-DEC-2001 (first entry)	
XX		
DE	Breast cancer-associated membrane protein (BCMP) known peptide #45.	
XX		
KW	Breast cancer-associated membrane protein; BCMP; cytosolic;	
XX		
OS	gene therapy.	
XX		
PN	Homo sapiens.	
XX	WO200162784-A2.	
PD	30-AUG-2001.	
XX		
PF	21-FEB-2001; 2001WO-GB00748.	
XX		
PR	25-FEB-2000; 2000GB-0004576.	
XX		
PR	21-DEC-2000; 2000GB-0031341.	
XX		
PA	(OXFO-) OXFORD GLYCOSCIENCES UK LTD.	
XX		
PI	Boyd RS, Stamps AC, Terrett JA;	
XX		
DR	WPI, 2001-557697/62.	
XX		
PT	Screening, diagnosis, prognosis, prevention and treatment of breast cancer in a subject, by using breast cancer-associated membrane	

PT proteins identified in membrane protein extracts of cultured human
 PT mammary cell lines -
 XX
 XX
 PS Claim 1; Page 16; 136pp; English.
 CC Sequences AAU25727-AAU25814 represent breast cancer-associated membrane
 CC proteins (BCMPs) and BCMP known and conceptual peptides. The BCMPs and
 CC their associated polynucleotides are used for treating or preventing
 CC breast cancer and also in prophylaxis and screening for breast cancer.
 CC The presence of or a predisposition to the disease can be detected by
 CC screening for BCMP expression and/or activity since the peptides may
 CC induce an immune response. The peptides are also useful for monitoring
 CC the effectiveness of breast cancer therapy.
 XX
 SO Sequence 10 AA;
 Query Match 100.0%; Score 22; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 00YQ 4
 3 00YQ 6
 Db 3 00YQ 6
 RESULT 2
 AAU10722
 ID AAU10722 standard; peptide; 13 AA.
 AC AAU10722;
 XX
 XX 09-APR-1997 (first entry)
 DT
 DE Catalytic light chain CDR3 peptide from phagemid library.
 XX
 XX CDR; complementarity determining region; Fab; fragment antigen;
 KM binding; antibody; identification; catalytically active; conjugate;
 KM immobilised suicide substrate.
 XX
 OS Synthetic.
 XX
 PN US5571681-A.
 XX
 PD 05-NOV-1996.
 XX
 PF 10-MAR-1994; 94US-0209525.
 XX
 XX 10-MAR-1994; 94US-0209525.
 (SCRI) SCRIPPS RES INST.
 XX
 PI Janda KD;
 XX
 DR WPI; 1996-505397/50.
 PT Identifying catalytically active proteins - using phage display
 PT protein library and immobilised suicide substrate for forming a
 PT covalent conjugate
 XX
 XX Example 7; Column 63-64; 50pp; English.
 PS
 CC AAU10720-W10725 are catalytically active light chain CDR3 peptides
 CC derived from a phagemid Fab display protein library. The library,
 CC which may contain light and heavy chain CDR3 peptides, are used with
 CC an immobilised suicide substrate (SS) to identify and obtain
 CC proteins/peptides with catalytic activity e.g. esterolytic, aldol
 CC condensation, beta-lactamase, glycosidase, RNase or proteolytic
 CC activity. The SS is an unreactive compound which has structural
 CC similarity to substrates of the CDR3 peptides and are activated by
 CC these peptides. The activation converts the SS to a species which,
 CC without prior release from the active site, binds covalently to the
 CC CDR3 peptide. The SS used for the CDR3 libraries was alpha-phenethyl
 CC pyridyl disulphide hapten and was coupled to bovine serum albumin.

XX
 SQ Sequence 13 AA;
 Query Match 100.0%; Score 22; DB 17; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 0Y 1 00YQ 4
 1 00YQ 4
 Db 1 00YQ 4
 RESULT 3
 AAU25184
 ID AAU25184 standard; peptide; 15 AA.
 AC AAU25184;
 XX
 XX 03-SEP-1999 (first entry)
 DT
 DE Foot and mouth disease peptide #3.
 XX
 XX Heat shock protein; HSP; complex; denatured protein matrix; antigen;
 KM vaccine; allergic disease; treatment; susceptibility; Th2; skin rash;
 KM allergic reaction; asthma.
 XX
 OS Foot and mouth disease virus.
 XX
 PN WO9929182-A1.
 XX
 XX 17-JUN-1999.
 PD
 PF 04-DEC-1998; 98WO-US25734.
 XX
 PF 05-DEC-1997; 97US-0986234.
 PR 05-DEC-1997; 97US-0985548.
 XX
 PA (UYNE-) UNIV NEW MEXICO STATE.
 XX
 PI Moseley PL, Wallen BS;
 XX
 DR WPI; 1999-394912/33.
 PT Synthesizing heat shock protein complexes using a denatured protein
 PT matrix
 XX
 PS Example 1; Fig 1a; 33pp; English.
 CC This invention describes a novel method for synthesizing heat shock
 CC protein (HSP) complexes comprising adding a heat shock protein to a
 CC denatured protein matrix for binding, and adding a complexing solution
 CC comprising a peptide to elute a heat shock protein-peptide complex. A
 CC HSP-antigen complex is useful as a vaccine for treating an allergic
 CC disease (in a mammal, preferably a human) to reduce susceptibility of
 CC the Th2 response, the complex comprising a HSP-antigenic peptide complex.
 CC The complex is administered to prevent a mammal from having an allergic
 CC reaction to an allergic disease, or administered to a mammal having an
 CC allergic disease, to reduce the allergic reactions. Allergic diseases
 CC include asthma and skin rashes. Prior art methods or preventing/treating
 CC allergic diseases include antihistamines which treat only the symptoms.
 CC corticosteroids which have severe side effects and desensitization
 CC therapy which has limited uses. The new method also allows more
 CC flexibility of use of peptide-based vaccines, as prior art HSP-based
 CC vaccines require isolation from a portion of the tumour itself. This
 CC sequence represents a peptide fragment from a foot and mouth disease
 CC protein which is used in the method of the invention.
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 22; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 XX ||||
 DB 9 QOYQ 12

RESULT 4
 ID AAY70487
 AAY70487 standard; Protein; 15 AA.

AC AAY70487;
 XX
 DT 04-JUL-2000 (first entry)

Synthesised peptide-2 from human prothrombin kringle-2 protein.

DE Human prothrombin kringle-2; prothrombin; kringle domain; cancer therapy;
 XX anti-endothelial cell proliferation; anti-tumour; angiogenesis;
 XX cancer; skin; laryngeal; uterine; colon; lung; bone marrow; cytostatic;
 KM anti-angiogenic.

Homo sapiens.

PN WO200014209-A1.

PD 16-MAR-2000.

PF 07-SEP-1999; 99WO-KR00525.

PR 07-SEP-1998; 98KR-0036786.

PA (KIMS/) KIM S. S.
 (GENO-) GENOTECH CORP.

PI Rhim TY, Park CS, Kim EK;

DR WPI; 2000-256965/22.

Human prothrombin kringle-1 and kringle-2 polypeptides that inhibit angiogenesis; useful for treating cancer -

Example 7; Page 19; 43pp; English.

CC The present sequence is a synthesised peptide of human prothrombin
 CC kringle-2 protein which is derived from human plasma prothrombin.
 CC Peptides were synthesised having amino acid sequences in order from
 CC N-terminal sequence of prothrombin kringle-2 by overlapping 2-5
 CC amino acids of both N-terminal and C-terminal of peptides. These peptides
 CC were purified by high performance liquid chromatography (HPLC). Kringle
 CC domains from prothrombin have anti-endothelial cell proliferation and
 CC anti-tumour activities. These inhibit angiogenesis which is involved in
 CC tumour growth. They are also useful for the treatment of cancer.
 CC e.g. skin, laryngeal, uterine, colon, lung and bone marrow cancer. The
 CC human prothrombin kringle-2 significantly suppresses
 CC tumour growth in vivo.

SQ Sequence 15 AA;

Query Match 100.0%; Score 22; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 XX ||||
 DB 11 QOYQ 14

RESULT 5
 ID AAY70488
 AAY70488 standard; Protein; 15 AA.

AC AAY70488;
 XX
 DT 04-JUL-2000 (first entry)

XX Synthesised peptide-3 from human prothrombin kringle-2 protein.

DE Human prothrombin kringle-2; prothrombin; kringle domain; cancer therapy;
 XX anti-endothelial cell proliferation; anti-tumour; angiogenesis;
 XX cancer; skin; laryngeal; uterine; colon; lung; bone marrow; cytostatic;
 KM anti-angiogenic.

Homo sapiens.

PN WO200014209-A1.

PD 16-MAR-2000.

PF 07-SEP-1999; 99WO-KR00525.

PR 07-SEP-1998; 98KR-0036786.

PA (KIMS/) KIM S. S.
 (GENO-) GENOTECH CORP.

PI Rhim TY, Park CS, Kim EK;

DR WPI; 2000-256965/22.

Human prothrombin kringle-1 and kringle-2 polypeptides that inhibit angiogenesis; useful for treating cancer -

Example 7; Page 19; 43pp; English.

CC The present sequence is a synthesised peptide of human prothrombin
 CC kringle-2 protein which is derived from human plasma prothrombin.
 CC Peptides were synthesised having amino acid sequences in order from
 CC N-terminal sequence of prothrombin kringle-2 by overlapping 2-5
 CC amino acids of both N-terminal and C-terminal of peptides. These peptides
 CC were purified by high performance liquid chromatography (HPLC). Kringle
 CC domains from prothrombin have anti-endothelial cell proliferation and
 CC anti-tumour activities. These inhibit angiogenesis which is involved in
 CC tumour growth. They are also useful for the treatment of cancer.
 CC e.g. skin, laryngeal, uterine, colon, lung and bone marrow cancer. The
 CC human prothrombin kringle-2 significantly suppresses
 CC tumour growth in vivo.

SQ Sequence 15 AA;

Query Match 100.0%; Score 22; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 XX ||||
 DB 1 QOYQ 4

RESULT 6
 ID ABG73642
 ABG73642 standard; Peptide; 17 AA.

AC ABG73642;

DT 06-MAR-2003 (first entry)

DE PERV antigenic peptide p15E.

KW PERV; porcine; env; gag; infection; vaccine; virucide; immunization;
 KW pig breeding; chromosome mapping; protection.

OS Porcine endogenous retrovirus.

PN DE10111433-A1.

PD 19-SEP-2002.

PF 09-MAR-2001; 2001DE-1011433.
XX
XX 09-MAR-2001; 2001DE-1011433.
BR
PA (BUND) BUNDESREPUBLIK DEUT PAUL-EHRICH-INST.
XX
PI Toenjes RR, Krach U;
XX WPI, 2002-699548/76.
DR
XX Replication-competent molecular clones of porcine endogenous
PT retrovirus, useful as sources of env peptides for vaccination and for
PT viral detection -
PS Example 11; Page 13, 52pp; German.

XX This invention describes novel replication-competent molecular clones of
CC porcine endogenous retrovirus (PERV) isolated from a porcine BAC
CC (bacterial artificial chromosome) library which are replication competent
CC after transfer into permissive cells. The invention also discloses (a)
XX env and gag polypeptides encoded by the molecular clones of the
XX invention; (b) porcine nucleic acids that represent the 5' and 3'
CC flanking sequences of the integration site of a replication competent
CC molecular clone in the porcine genome; (c) oligonucleotides for detecting
CC integrated PERV; (d) a method for detecting appearance of infectious PERV
CC particles in a sample; (e) a vaccine for immunization against
CC replication-competent PERV, comprising an env protein and (f) a method
CC for isolating the molecular clones of the invention. The products of the
CC invention are have virocid activity. The clones are from native PERV
CC and make possible comparisons of proviral sequences of different origins
CC at the molecular, structural and cellular levels, and mapping of proviral
CC sequences at chromosomal positions in specific breeds of pigs. Also
CC different breeds of pigs can be screened for presence of specific
CC infectious PERV, i.e., to identify PERV-free animals. Also vaccines that
CC contain the env polypeptides encoded by the clones of the invention are
CC useful for protection against replication-competent PERV. The products of
CC the invention are functional, replication-competent, full-length,
CC proviral clones isolated directly from the pig genome, i.e. they are
CC 'native' PERV. This sequence represents an antigenic peptide associated
CC with the PERV clones described in the disclosure of the invention.

SO Sequence 17 AA;

Query Match 100.0%; Score 22; DB 23; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.66+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 OOOO 4
|||
4 OOOO 7

RESULT 7
ABG75795 standard; Protein; 21 AA.
XX ABG75795;
AC
XX
DT 01-MAY-2003 (first entry)
XX
DB Bovine histactin like polypeptide, #5.
XX
XX Cow; histactin; bovine; salivary secretion; oral cavity;
KW non-immune defence system; oral candidosis; gene therapy;
KW microbial infection; fungal infection; dental caries; tartar;
KW cystic fibrosis; systemic infection; Candida infection; mastitis;
KW fungicide; antibacterial.
XX
XX Boe taurus.
OS
XX
XX US2002164625-A1.
PN
XX
PD 07-NOV-2002.

XX	19-FEB-2002; 2002US-0079754.
PF	
XX	23-AUG-1999; 99US-150330P.
PR	29-OCT-1999; 99US-162701P.
PR	22-AUG-2000; 2000US-0644190.
PR	27-OCT-2000; 2000US-0699146.
XX	
PA	(GENE-) GENESIS RES & DEV CORP. LTD.
PI	Glenn M, Grigor MR, Molenaar AJ, Davis SR;
XX	
DR	WPI; 2003-275306/27.
DR	N-PSDB; ABX11392.
XX	
PT	New histatin polynucleotides and polypeptides expressed in bovine
PT	tissues, useful for treating bacterial or fungal infections in mammals,
PT	e.g. infections of the oral cavity, vagina, urethra, ear or skin, or
PT	systemic infections -
XX	
PS	Claim 6; Page 14; 17pp; English.
XX	
CC	The invention discloses isolated histatin polynucleotides and
CC	polypeptides, which comprise sequences expressed in bovine. Histatins are
CC	histidine rich and specific to the salivary secretions. They are believed
CC	to function as part of the non-immune defence system, particularly in the
CC	oral cavity and have promise as therapeutic agents in humans with oral
CC	candidosis. Also disclosed are compositions comprising the histatin
CC	polypeptide or polynucleotide, and at least one component consisting of
CC	physiological or pharmaceutical carriers or immunostimulants. The
CC	histatin polynucleotide, polypeptide, or the cosmetic composition
CC	comprising the polypeptide, is useful for treating (e.g. gene therapy) a
CC	disorder in a mammal, particularly microbial or fungal infections. The
CC	histatin polynucleotide is also useful in genome mapping, physical
CC	mapping or in the positional cloning of genes. Specifically, the
CC	polypeptide or polynucleotide is useful for treating fungal or bacterial
CC	infections of the oral cavity (e.g. dental caries, plaque or tartar),
CC	vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis
CC	caused by Pseudomonas aeruginosa), mucosa or eye, as well as for treating
CC	systemic infections such as systemic Candida infection. The polypeptide
CC	or polynucleotide is also useful for veterinary applications e.g. for
CC	treating mastitis. The sequence presented is a bovine histatin like
CC	polypeptide.
XX	
XX	Sequence 21 AA;
SQ	
Query Match	100.0%; Score 22; DB 24; Length 21;
Best Local Similarity	100.0%; Pred. No. 2e+02;
Matches 4; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1 QOYQ 4
DB	2 QOYQ 5
RESULT 8	
AA007536	
ID	AA007536 standard; Protein; 32 AA.
AC	
XX	AA007536;
DT	06-NOV-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 21428.
XX	
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;
XX	nervous system disorders; arthritis; inflammation.
OS	Homo sapiens.
PN	WO200164835-A2.

XX 07-SEP-2001.
 PD 26-FEB-2001; 2001WO-US04927.
 XX 28-FEB-2000; 2000US-0515126.
 PF 18-MAY-2000; 2000US-0577409.
 PR (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI, 2001-514838/56.
 DR N-PSDB; AA187467.
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX Claim 20; SEQ ID NO 21428; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AA179941-AA193841) and the encoded proteins (AA000010-AA013910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activity/inhibit activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 32 AA;

Query Match 100.0%; Score 22; DB 22; Length 32;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 DB 12 QOYQ 15

RESULT 9

AAE32369 standard; peptide; 32 AA.

AAE32369;

24-MAR-2003 (first entry)

Human HBNF peptide #5.

Vascular endothelial growth factor; VEGF; angiogenesis; wound healing;
 bone growth; osteoporosis; osteoarthritis; bone reconstruction; ulcer;
 lesion; injury; trauma; periodontal condition; protein therapy; human;
 HBNF.

Homo sapiens.

WO200283851-A2.

24-OCT-2002.

10-APR-2002; 2002WO-US11406.

10-APR-2001; 2001US-0832355.

(GENV-) GENVEC INC.

XX Kovesdi I, Kessler PD;
 PF WPI; 2003-075536/07.
 DR New fusion protein comprising a non-heparin-binding vascular
 PT endothelial growth factor (VEGF) peptide portion and a non-VEGF peptide
 PT portion, useful for promoting angiogenesis and/or bone growth in
 PT mammals -

Dislosure; Page 140; 191pp; English.

The invention relates to a fusion protein comprising non-heparin binding vascular endothelial growth factor (VEGF) peptide portion and a non-VEGF peptide portion useful for promoting angiogenesis and/or bone growth in mammalian host. The fusion protein is useful for promoting angiogenesis, wound healing and bone growth. Compositions containing bone growth promoting fusion protein can be used to treat osteoporosis, rheumatoid or osteoarthritis, to improve poor bone healing, to promote implant integration and function of artificial joints and to facilitate bone reconstruction. They can also be used to treat e.g. ulcers, lesions, injuries, burns, trauma, periodontal conditions, lacerations and other conditions. The invention is also useful in protein therapy. The present sequence is human HBNF peptide. This sequence is used in the invention.

XX Sequence 32 AA;

Query Match 100.0%; Score 22; DB 24; Length 32;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 DB 4 QOYQ 7

RESULT 10

AAE47591 standard; Protein; 42 AA.

AAE47591;

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 60000.

Protein identification; signal transduction pathway; metabolic pathway;
 hybridisation assay; genetic mapping; gene expression control; promoter;
 termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.

05-MAR-1999; 99US-0123180.

09-MAR-1999; 99US-0123548.

23-MAR-1999; 99US-0125788.

25-MAR-1999; 99US-0126264.

29-MAR-1999; 99US-0126785.

01-APR-1999; 99US-0127462.

06-APR-1999; 99US-0128234.

08-APR-1999; 99US-0128714.

16-APR-1999; 99US-0129845.

19-APR-1999; 99US-0130077.

21-APR-1999; 99US-0130449.

23-APR-1999; 99US-0130510.

23-APR-1999; 99US-0130891.

28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0133484.
PR 05-MAY-1999; 99US-0133485.
PR 06-MAY-1999; 99US-0133486.
PR 07-MAY-1999; 99US-0133487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134377.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136352.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148668.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR	11-MAY-1999;	99US-0134256;
PR	14-MAY-1999;	99US-01343218;
PR	14-MAY-1999;	99US-01343219;
PR	14-MAY-1999;	99US-01344221;
PR	14-MAY-1999;	99US-0134370;
PR	14-MAY-1999;	99US-0134766;
PR	18-MAY-1999;	99US-0134941;
PR	19-MAY-1999;	99US-0135121;
PR	21-MAY-1999;	99US-0135353;
PR	21-MAY-1999;	99US-0135629;
PR	24-MAY-1999;	99US-0136021;
PR	25-MAY-1999;	99US-0136392;
PR	28-MAY-1999;	99US-0136782;
PR	01-JUN-1999;	99US-0137222;
PR	03-JUN-1999;	99US-0137528;
PR	04-JUN-1999;	99US-0137502;
PR	07-JUN-1999;	99US-0137724;
PR	08-JUN-1999;	99US-0138094;
PR	10-JUN-1999;	99US-0138540;

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PK 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.

PR 14-JUL-1999: 99US-0143624.
PR 13-JUL-1999: 99US-0143542.
PR 12-JUL-1999: 99US-0142577.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 13-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.

PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	22-JUL-1999;	99US-0145192.
PR	22-JUL-1999;	99US-0145192.

PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145593.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147320.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 10-AUG-1999; 99US-0147935.
 PR 11-AUG-1999; 99US-0148171.
 PR 12-AUG-1999; 99US-0148319.
 PR 13-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148585.
 PR 16-AUG-1999; 99US-0148684.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151085.
 PR 27-AUG-1999; 99US-0151086.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 22; DB 21; Length 45;
 Best Local Similarity 100.0%; Pred. No. 4,3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 21 QOYQ 24

RESULT 12
 AAY26934
 ID AAY26934 standard; peptide; 46 AA.
 XX

AC AAY26934;

DT 21-DEC-1999 (first entry)

DE Human herpesvirus 6 recombinase D35 motif.

KW Recombinase motif; retrovirus; integrase; invertebrate; transposase;
 KW consensus; vaccine; gene therapy; allergy; autoimmune disease; cancer;
 KW cardiovascular disease; graft rejection; hematopoietic disorder;
 KW immunosuppressive disorder; immunoproliferative disease; jaundice;
 KW immunodeficiency disease; infectious disease; inflammatory disease;
 KW septic shock; metabolic defect.

OS Human herpesvirus 6.

XX US5959074-A.

PN 28-SEP-1999.

PD 28-FEB-1997; 97US-0807332.

PP 01-MAR-1996; 96US-0012616.

PR 02-AUG-1996; 96US-0023064.

PA (NAME-) NAT JEWISH MEDICAL & RES CENT.

XX Gelfand EW, Dreyfus DH;

PI WPI; 1999-561074/47.

DR Peptides derived from a Herpes virus recombinogenic motif useful for the

PT production of vaccines and in gene therapy -

XX Claim 1; Column 37-38; 29pp; English.

CC The invention relates to novel recombinogenic motifs found in
 CC (retro)viral integrases which have similarity to invertebrate
 CC transposase molecules. The peptides, or their homologues, have the
 CC following identifying characteristics: (a) an initial and terminal amino
 CC acid comprising Asp or Glu, separated by 34-39 amino acids, especially
 CC 35 amino acids (the D35 motif); (b) a high probability (p < 0.05) of
 CC alignment with a consensus sequence (AAY26941), as determined by:
 CC (1) aligning the initial and terminal residues of the homologue sequence
 CC with those of the consensus; (2) aligning the intervening residues of
 CC the homologue with those of the consensus by maintaining the spacing of
 CC the consensus while, if necessary, altering (by the insertion of spaces
 CC or the deletion of residues) the spacing of the homologue; and

(3) generating a distribution of 1500 random sequences of amino acids between the initial and terminal residues and identifying the percentage of aligned residues. The recombinogenic motifs may be useful in the production of vaccines and in gene therapy to treat allergies, CC autoimmune diseases, cancers, cardiovascular diseases, graft rejection, CC hematopoietic disorders, immunosuppressive disorders, CC immunoproliferative diseases, immunodeficiency diseases, infectious CC diseases, inflammatory diseases, jaundice, septic shock and other CC immunological, genetic or metabolic defects.

Sequence 46 AA;

Query Match 100.0%; Score 22; DB 20; Length 46;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
15 QOYO 18

RESULT 13

ID AAB73930 standard; protein; 46 AA.

AC AAB73930;

DT 29-MAY-2001 (first entry)

DE Human herpesvirus 6 DBP recombinogenic sequence.

Human herpesvirus 6; HH6; D35E; anti-viral; immunosuppressant;
immunostimulant; cytotoxic; anti-allergic; cardiant; anti-inflammation;
gene therapy; infection; herpes virus replication; V(D)J recombination;
retroviral integrase; immunoglobulin class switching; allergy;
autoimmune disease; cancer; cardiovascular disease; graft rejection;
haematopoietic disorder; immunological disorder; DNA binding protein;
DBP.

OS Human herpesvirus 6.

PN US6187584-B1.

PD 13-FEB-2001.

PF 23-JUN-1999; 99US-0338876.

XX 01-MAR-1996; 96US-0012616.

XX 02-AUG-1996; 96US-0023064.

XX 28-FEB-1997; 97US-0807332.

PA (NABJ-) NAT JEWISH MEDICAL & RBS CENT.

PI Dreyfus DH, Gelfand EW;

DR WPI; 2001-256131/26.

PT New nucleic acid molecules encoding peptides with transposase
activities, useful as immunosuppressants, anti-viral agents, or vectors
for gene therapy of diseases, e.g. allergies, autoimmune diseases,
cancers or graft rejection

PS Claim 1; Column 25; 31pp; English.

XX The present sequence is provided in a specification relating to a new
isolated nucleic acid molecule comprising a nucleic acid sequence
encoding a peptide consisting of one of ten amino acid sequences
defined in the specification or homologues of these sequences.

CC The nucleic acid molecules are useful for producing peptides
important in the regulation and function of herpes virus replication,
V(D)J recombination, retroviral integrase function or immunoglobulin
class switching. The isolated nucleic acid molecules and the peptides
they encode are useful as immunosuppressants, anti-viral agents or

CC vectors for gene therapy. These are particularly useful in protecting
humans from a variety of diseases, e.g. allergies, autoimmune disease,
cancers, cardiovascular diseases, graft rejection, haematopoietic
disorders, immunodeficiency diseases, immunoproliferative diseases,
CC immunosuppressive disorders, infectious diseases, inflammatory diseases,
CC jaundice, septic shock, or other immunological defects, as well as other
genetic or metabolic defects.

Sequence 46 AA;

Query Match 100.0%; Score 22; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
15 QOYO 18

RESULT 14

ID AAG47589 standard; Protein; 49 AA.

AC AAG47589;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 59998.

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134968.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149366.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match Score 22; DB 21; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||
Db 25 00YQ 28

RESULT 15
ABU67216

ID ABU67216 standard; Protein; 51 AA.

XX AC ABU67216;

DT 28-MAY-2003 (first entry)

G-protein coupled receptor (GPCR) fragment #15.

KM Human; G-protein coupled receptor; receptor; GPCR; palmitoylation site;
KM phosphorylation cluster; arrestin; endosome; angina pectoris; rhinitis;
KM atherosclerosis; asthma; emphysema; inflammatory disease; glaucoma; pain;
KM rheumatoid arthritis; obesity; Parkinson's disease; rat.

XX OS Homo sapiens.

FN US2002106739-A1.

PD 08-AUG-2002.

PF 05-NOV-2001; 2001US-0993844.

PR 03-NOV-2000; 2000US-245772P.

PR 08-JAN-2001; 2001US-260363P.

XX (OAKLEY R. H.

PA (BARAK L. S.

PA (LAPORTE S. A.

PA (CARON M. G.

PI OAKLEY RH, BARAK LS, Laporte SA, Caron MG;

DR WPI; 2002-690758/74.

XX Modified G-protein coupled receptor useful for identifying an agonist,
XX inverse agonist or antagonist of the receptor, comprises a carboxyl
XX terminal having one or more clusters of phosphorylation -

PS Disclosure; Fig 2; 57pp; English.

CC The invention relates to a modified G-protein coupled receptor (GPCR)
CC comprising an NPXY motif, and a carboxyl terminal tail which comprises a
CC putative site of palmitoylation and clusters of phosphorylation, and a
CC retained portion of a carboxyl terminal region of a GPCR portion fused to
CC a portion of the carboxyl terminal from a second GPCR, that comprises
CC phosphorylation clusters and a putative palmitoylation site 10-25 amino
CC acid residues downstream of a second NPXY motif. The modified GPCR is
CC useful for screening compounds for GPCR activity which comprises
CC providing a cell that expresses at least one modified GPCR, where the
CC cell further comprises arrestin conjugated to a detectable molecule,
CC exposing the cell to the compound, detecting the location of the arrestin
CC within the cell, comparing the location of the arrestin within the cell
CC in the presence of the compound to the location of the arrestin within
CC the cell in the absence of the compound and correlating a difference
CC between the location of arrestin within the cell in the presence of the
CC compound and the presence of the location of the arrestin within the cell
CC in the absence of the compound. Preferably, the arrestin is detected in
CC endosomes. The GPCR and a nucleic acid encoding the modified GPCR are
CC useful for preventing and/or treating a disease associated with GPCR in
CC mammals, such as angina pectoris, atherosclerosis, asthma, emphysema,

CC rhinitis, inflammatory disease, rheumatoid arthritis, glaucoma, pain,
CC obesity or Parkinson's disease, by modulating GPCR activity and affinity
CC for arrestin. Sequences ABU67202-ABU67240 represent GPCR polypeptide
CC fragments used in the scope of the invention.

XX Sequence 51 AA;

QY 1 00YQ 4
|||
Db 32 00YQ 35

RESULT 16
ABP53997

ID ABP53997 standard; peptide; 51 AA.

XX AC ABP53997;

DT 10-JAN-2003 (first entry)

Human muscarinic acetylcholine receptor M3 class A GPCR.

KM G-protein-coupled receptor; GPCR; detection; receptor; olfaction;
KM photo-transduction; neurotransmission; vascular tone; cardiac output;
KM digestion; pain; fluid electrolyte balance.

XX OS Homo sapiens.

FN WO200273199-A1.

PD 19-SEP-2002.

PF 13-MAR-2002; 2002WO-US07434.

PR 13-MAR-2001; 2001US-275339P.

PR 12-MAR-2002; 2002US-0095620.

XX (UYDU-) UNIV DUKE.

PA BARAK LS, OAKLEY RH;

PI WPI; 2003-018747/01.

DR

XX Detecting G-protein-coupled receptor pathway activity comprises

PT treating a cell having G-protein-coupled receptor activity with a test

PT compound, obtaining a digital image of the cell and detecting position

PT of conjugated proteins -

XX Disclosure; Page 26; 42pp; English.

CC The present invention describes a method for detecting G-protein-coupled
CC receptor (GPCR) pathway activity. The method comprises: (1) providing at
CC least one cell expressing GPCR and conjugated proteins; (2) treating the
CC cell with a test compound; (3) obtaining a digital image of the cell by
CC detecting and measuring energy emitted from the detectable molecules; and
CC (4) detecting the localization of some of the conjugated proteins at one
CC of endocytic vesicles and endosomes. The method can be used for detecting
CC compounds that activate GPCR-activity in vivo and in vitro, where the
CC compound is useful for regulating various physiological functions
CC including photo-transduction, olfaction, neurotransmission, vascular
CC tone, cardiac output, digestion, pain and fluid electrolyte balance.
CC The present sequence represents a GPCR amino acid sequence which is given
CC in the exemplification of the present invention.

XX Sequence 51 AA;

QY 1 00YQ 4
|||
Db 32 00YQ 35

RESULT 16
ABP53997
ID ABP53997 standard; peptide; 51 AA.
XX AC ABP53997;
DT 10-JAN-2003 (first entry)
Human muscarinic acetylcholine receptor M3 class A GPCR.
KM G-protein-coupled receptor; GPCR; detection; receptor; olfaction;
KM photo-transduction; neurotransmission; vascular tone; cardiac output;
KM digestion; pain; fluid electrolyte balance.
XX OS Homo sapiens.
FN WO200273199-A1.
PD 19-SEP-2002.
PF 13-MAR-2002; 2002WO-US07434.
PR 13-MAR-2001; 2001US-275339P.
PR 12-MAR-2002; 2002US-0095620.
XX (UYDU-) UNIV DUKE.
PA BARAK LS, OAKLEY RH;
PI WPI; 2003-018747/01.
DR
XX Detecting G-protein-coupled receptor pathway activity comprises
PT treating a cell having G-protein-coupled receptor activity with a test
PT compound, obtaining a digital image of the cell and detecting position
PT of conjugated proteins -
XX Disclosure; Page 26; 42pp; English.
CC The present invention describes a method for detecting G-protein-coupled
CC receptor (GPCR) pathway activity. The method comprises: (1) providing at
CC least one cell expressing GPCR and conjugated proteins; (2) treating the
CC cell with a test compound; (3) obtaining a digital image of the cell by
CC detecting and measuring energy emitted from the detectable molecules; and
CC (4) detecting the localization of some of the conjugated proteins at one
CC of endocytic vesicles and endosomes. The method can be used for detecting
CC compounds that activate GPCR-activity in vivo and in vitro, where the
CC compound is useful for regulating various physiological functions
CC including photo-transduction, olfaction, neurotransmission, vascular
CC tone, cardiac output, digestion, pain and fluid electrolyte balance.
CC The present sequence represents a GPCR amino acid sequence which is given
CC in the exemplification of the present invention.
XX Sequence 51 AA;
QY 1 00YQ 4
|||
Db 32 00YQ 35

QY 1 QOYQ 4
DB 32 QOYQ 35

RESULT 17

AA084018
ID AA084018 standard; Protein; 55 AA.

AC AA084018;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen SEQ ID NO:11611.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW Cytostatic; gene therapy; vaccine; metastasis.

XX Homo sapiens.

MO200157182-A2.

PD 09-AUG-2001.

PF 17-JAN-2001; 2001MO-US01354.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225477.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0225779.

PR 22-AUG-2000; 2000US-0226811.

PR 22-AUG-2000; 2000US-0226868.

PR 23-AUG-2000; 2000US-0227182.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235835.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 02-OCT-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239335.

PR 13-OCT-2000; 2000US-0239337.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 01-NOV-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.

PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.

PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

N-PSDB; AAK56799.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Claim 11; SEQ ID NO 11611; 3071bp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAK67694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention.

Sequence 55 AA;

Query Match 100.0%; Score 22; DB 22; Length 55;

Best Local Similarity 100.0%; Pred. No. 5.2e+02; Mismatches 0; Indels 0; Gaps 0;

OY 1 OQYO 4
 ||||
 Db 11 OQYO 14

RESULT 18

AAK12599

ID AAG12599 standard; Protein; 58 AA.

XX AAG12599;

DT 17-OCT-2000 (first entry)

DE Zea mays protein fragment SEQ ID NO: 11775.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.

XX Zea mays subsp. mays.

PN EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000BP-0301439.

XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150584.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 22; DB 21; Length 58;
Best Local Similarity 100.0%; Pred. No. 5; Se+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQYQ 4
|||||
Db 36 QQYQ 39

RESULT 19
AAG41160
ID AAG41160 standard; Protein; 58 AA.
XX
AC AAG41160;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 51174.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX

PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149425.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161559.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 22; DB 21; Length 58;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYO 4
 ||||
 36 QOYO 39

RESULT 20

ABG75791 standard; Protein; 58 AA.

AC ABG75791;
 XX
 DT 01-MAY-2003 (first entry)
 XX
 DE Bovine histatin like polypeptide, #1.
 XX
 KW Cow; histatin; bovine; salivary secretion; oral cavity;
 KW non-immune defence system; oral candidosis; gene therapy;
 KW microbial infection; fungal infection; dental caries; plaque; tartar;
 KW cystic fibrosis; systemic infection; Candida infection; mastitis;
 XX
 OS Bos taurus.
 XX
 PN US2002164625-A1.
 XX
 PD 07-NOV-2002.
 XX
 PF 19-FEB-2002; 2002US-0079754.
 XX
 XX 23-AUG-1999; 99US-150330P.
 PR 29-OCT-1999; 99US-162701P.
 PR 22-AUG-2000; 2000US-0644190.

PR 27-OCT-2000; 2000US-0699146.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Glenn M, Grigor MR, Molenaar AJ, Davis SR,
 XX
 DR WPI; 2003-275306/27.
 DR N-PSDB; ABX1388.
 XX
 PT New histatin polynucleotides and polypeptides expressed in bovine
 PT tissues, useful for treating bacterial or fungal infections in mammals,
 PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or
 XX
 PS Claim 6; Page 13; 17pp; English.

CC The invention discloses isolated histatin polynucleotides and
 CC polypeptides, which comprises sequences expressed in bovine. Histatins are
 CC histidine rich and specific to the salivary secretions. They are believed
 CC to function as part of the non-immune defence system, particularly in the
 CC oral cavity and have promise as therapeutic agents in humans with oral
 CC candidosis. Also disclosed are compositions comprising the histatin
 CC polypeptide or polynucleotide, and at least one component consisting of
 CC physiological or pharmaceutical carriers or immunostimulants. The
 CC histatin polynucleotide, polypeptide, or the cosmetic composition
 CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a
 CC disorder in a mammal, particularly microbial or fungal infections. The
 CC histatin polynucleotide is also useful in genome mapping, physical
 CC mapping or in the positional cloning of genes. Specifically, the
 CC polypeptide or polynucleotide is useful for treating fungal or bacterial
 CC infections of the oral cavity (e.g. dental caries, plaque or tartar),
 CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis
 CC caused by *Pseudomonas aeruginosa*), mucosa or eye, as well as for treating
 CC systemic infections such as systemic Candida infection. The polypeptide
 CC or polynucleotide is also useful for veterinary applications e.g. for
 CC treating mastitis. The sequence presented is a bovine histatin like
 CC polypeptide.

SO Sequence 58 AA;

Query Match 100.0%; Score 22; DB 24; Length 58;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYO 4
 ||||
 39 QOYO 42

RESULT 21

ABG75792 standard; Protein; 58 AA.

AC ABG75792;
 XX
 DT 01-MAY-2003 (first entry)
 XX
 DE Bovine histatin like polypeptide, #2.
 XX
 KW Cow; histatin; bovine; salivary secretion; oral cavity;
 KW non-immune defence system; oral candidosis; gene therapy;
 KW microbial infection; fungal infection; dental caries; plaque; tartar;
 KW cystic fibrosis; systemic infection; Candida infection; mastitis;
 KW fungicide; antibacterial.
 XX
 OS Bos taurus.
 XX
 PN US2002164625-A1.
 XX
 PD 07-NOV-2002.
 XX
 PF 19-FEB-2002; 2002US-0079754.
 XX

PR 23-AUG-1999; 99US-150330P.
 PR 29-OCT-1999; 99US-162701P.
 PR 22-AUG-2000; 2000US-0644190.
 PR 27-OCT-2000; 2000US-0699146.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Glenn M, Grigor MR, Molenaar AJ, Davis SR;
 XX
 DR WPI; 2003-275306/27.
 DR N-PSDB; ABX11389.
 XX
 PT New histatin polynucleotides and polypeptides expressed in bovine
 PT tissues, useful for treating bacterial or fungal infections in mammals,
 PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or
 PT systemic infections -
 XX
 PS Claim 6, Page 13, 17pp; English.
 XX

The invention discloses isolated histatin polynucleotides and polypeptides, which comprise sequences expressed in bovine. Histatins are histidine rich and specific to the salivary secretions. They are believed to function as part of the non-immune defence system, particularly in the oral cavity and have promise as therapeutic agents in humans with oral candidosis. Also disclosed are compositions comprising the histatin polypeptide or polynucleotide, and at least one component consisting of physiological or pharmaceutical carriers or immunostimulants. The histatin polynucleotide, polypeptide, or the cosmetic composition comprising the polypeptide, is useful for treating (e.g. gene therapy) a disorder in a mammal, particularly microbial or fungal infections. The histatin polynucleotide is also useful in genome mapping, physical mapping or in the positional cloning of genes. Specifically, the polypeptide or polynucleotide is useful for treating fungal or bacterial infections of the oral cavity (e.g. dental caries, plaque or tartar), vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis caused by *Pseudomonas aeruginosa*), mucosa or eye, as well as for treating systemic infections such as systemic *Candida* infection. The polypeptide or polynucleotide is also useful for veterinary applications e.g. for treating mastitis. The sequence presented is a bovine histatin like polypeptide.

Sequence 58 AA;

Query Match 100.0%; Score 22; DB 24; Length 58;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYO 4
 ||||
 39 QOYO 42

RESULT 22

ABG75794
 ID ABG75794 standard; Protein; 58 AA.

XX
 AC ABG75794;
 XX

DT 01-MAY-2003 (first entry)
 XX

DE Bovine histatin like polypeptide, #4.
 XX

XX Cow; histatin; bovine; salivary secretion; oral cavity;
 XX non-immune defence system; oral candidosis; gene therapy;
 XX microbial infection; fungal infection; dental caries; plaque; tartar;
 XX cystic fibrosis; systemic infection; *Candida* infection; mastitis;
 XX fungicide; antibacterial.
 XX

OS Bos taurus.
 XX

PN US2002164625-A1.
 XX

PD 07-NOV-2002.
 XX

XX
 PF 19-FEB-2002; 2002US-0079754.
 XX
 PR 23-AUG-1999; 99US-150330P.
 PR 29-OCT-1999; 99US-162701P.
 PR 22-AUG-2000; 2000US-0644190.
 PR 27-OCT-2000; 2000US-0699146.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Glenn M, Grigor MR, Molenaar AJ, Davis SR;
 XX
 DR WPI; 2003-275306/27.
 DR N-PSDB; ABX11391.
 XX
 PT New histatin polynucleotides and polypeptides expressed in bovine
 PT tissues, useful for treating bacterial or fungal infections in mammals,
 PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or
 PT systemic infections -
 XX
 PS Claim 6, Page 14, 17pp; English.
 XX

The invention discloses isolated histatin polynucleotides and polypeptides, which comprise sequences expressed in bovine. Histatins are histidine rich and specific to the salivary secretions. They are believed to function as part of the non-immune defence system, particularly in the oral cavity and have promise as therapeutic agents in humans with oral candidosis. Also disclosed are compositions comprising the histatin polypeptide or polynucleotide, and at least one component consisting of physiological or pharmaceutical carriers or immunostimulants. The histatin polynucleotide, polypeptide, or the cosmetic composition comprising the polypeptide, is useful for treating (e.g. gene therapy) a disorder in a mammal, particularly microbial or fungal infections. The histatin polynucleotide is also useful in genome mapping, physical mapping or in the positional cloning of genes. Specifically, the polypeptide or polynucleotide is useful for treating fungal or bacterial infections of the oral cavity (e.g. dental caries, plaque or tartar), vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis caused by *Pseudomonas aeruginosa*), mucosa or eye, as well as for treating systemic infections such as systemic *Candida* infection. The polypeptide or polynucleotide is also useful for veterinary applications e.g. for treating mastitis. The sequence presented is a bovine histatin like polypeptide.

Sequence 58 AA;

Query Match 100.0%; Score 22; DB 24; Length 58;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 ||||
 Db 39 QOYO 42

RESULT 23

AAG15832
 ID AAG15832 standard; Protein; 60 AA.

XX
 AC AAG15832;
 XX

DT 17-OCT-2000 (first entry)
 XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 16239.
 XX

XX Protein identification; signal transduction pathway; metabolic pathway;
 XX hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 XX

OS Arabidopsis thaliana.
 XX

PN EP1033405-A2.
 XX

PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140553.
PR 23-JUN-1999; 99US-0140554.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149126.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154033.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.


```

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 100.0%; Score 22; DB 21; Length 60;

Best Local Similarity 100.0%; Pred. No. 5.7e+02; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 00YQ 4
    |||
Db 35 00YQ 38

```

ULT 24

ID A6G52856 standard; Peptide; 64 AA.

XX A6G52856;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID No 31504.

XX Human, liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00664.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

```

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
PR (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver
XX Claim 27; SEQ ID No 31504; 658bp; English.
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver (I) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABG47348-ABG59930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 64 AA;

```

Query Match 100.0%; Score 22; DB 22; Length 64;

Best Local Similarity 100.0%; Pred. No. 6.1e+02; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 00YQ 4
    |||
Db 40 00YQ 43

```

RESULT 25

ID ABB38011 standard; Peptide; 64 AA.

XX ABB38011;

DT 04-FEB-2002 (first entry)

DE Peptide #5517 encoded by human foetal liver single exon probe.

XX Human, foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 PS Claim 27; SEQ ID NO 30646; 639pp + sequence listing; English.
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 64 AA;
 Query Match 100.0%; Score 22; DB 22; Length 64;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYQ 4
 DB 40 QOYQ 43
 RESULT 26
 AAM58638
 ID AAM58638 standard; Protein; 64 AA.
 XX
 AC AAM58638;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30743.
 XX
 DE Human; brain expressed exon; gene expression analysis; probe;
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KM epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 OS WO200157275-A2.
 PN
 PD 09-AUG-2001.
 XX
 PD 30-JAN-2001; 2001WO-US00667.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2001-483446/52.
 DR
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 30743; 650pp + Sequence listing; English.
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 SQ Sequence 64 AA;
 Query Match 100.0%; Score 22; DB 22; Length 64;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYQ 4
 DB 40 QOYQ 43
 RESULT 27
 AAM71144
 ID AAM71144 standard; Protein; 64 AA.
 XX
 AC AAM71144;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31450.
 XX
 DE Human; bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; Leukaemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 OS WO200157276-A2.
 PN
 PD 09-AUG-2001.
 XX
 PD 30-JAN-2001; 2001WO-US00668.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2001-483900/53.
 DR
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 31450; 658pp + Sequence listing; English.
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 SQ Sequence 64 AA;
 Query Match 100.0%; Score 22; DB 22; Length 64;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYQ 4
 DB 40 QOYQ 43

RESULT 28
 AAM31425
 ID AAM31425 standard; Protein; 64 AA.
 XX
 AC AAM31425;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Peptide #5462 encoded by probe for measuring placental gene expression.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PS 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 27; SEQ ID No 31694; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP;
 CC see AAI31315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 CC
 XX Sequence 64 AA;
 XX
 Query Match 100.0%; Score 22; DB 22; Length 64;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 00YQ 4
 ||||
 DB 40 00YQ 43
 XX
 RESULT 29
 ABG40941
 ID ABG40941 standard; Peptide; 64 AA.
 XX
 AC ABG40941;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 30606.
 XX
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberos scleriosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PS 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 27; SEQ ID No 30606; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberos scleriosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SEQ Sequence 64 AA; Query Match 100.0%; Score 22; DB 23; Length 64;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
40 QOYO 43
Db 40 QOYO 43

RESULT 30
ABP64187
ID ABP64187 standard; Protein; 66 AA.
XX
AC ABP64187;
XX
DT 04-NOV-2002 (first entry)
XX
XX Human ORF557.

Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
cancer; cardiovascular disease; allergy; autoimmune disease;
wound healing; blood coagulation disorder; inflammatory disorder.

OS Homo sapiens.
XX
PN US2002082206-A1.
XX
PD 27-JUN-2002.
XX
PF 30-MAY-2001; 2001US-0867550.
XX
PR 30-MAY-2000; 2000US-208427P.
XX
XX (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.

Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
MPI: 2002-626554/67.
DR N-PSDB; ABQ98750.

New polypeptide designated ORFX are present in human atherogenic cells
and are useful to prevent and treat ORFX-associated disorders including
inflammatory disease -
cancer, allergy, wound healing or autoimmune, cardiovascular or
inflammatory disease -
Claim 10; SEQ ID 1114; 78pp; English.

The present invention relates to novel human ORFX polypeptides and their
coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
were discovered in human atherogenic cells, in particular in platelets
and human umbilical vein endothelial cells (HUVEC) and are expressed in
many other tissues as well. Atherogenic cells are cells which have the
potential to develop atherosclerotic plaques. The ORFX polypeptides and
nucleic acids are useful for treating or preventing a pathological
condition associated with an ORFX-associated disorder, e.g. cancer,
cardiovascular disease, allergy, autoimmune disease, wound healing, blood
coagulation disorders or inflammatory disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site at seqdata.uspto.gov/sequence.html?DocID=20020082206.

XX
SQ Sequence 66 AA;
Query Match 100.0%; Score 22; DB 23; Length 66;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
16 QOYO 19
Db 16 QOYO 19

RESULT 31
AAM88630
ID AAM88630 standard; Protein; 68 AA.
XX
AC AAM88630;
XX
DT 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen SEQ ID NO:16223.
DE Human immune/haematopoietic antigen SEQ ID NO:16223.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytosstatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228927.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229345.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.

PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249287.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX	(HUMA-) HUMAN GENOME SCT INC.	
PA		
XX	Rosen CA, Barash SC, Ruben SM;	
P1		
DR	WPI; 2001-463426/52.	
XX	N-PSDB; AAK61411.	
PT		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
PT	metastasis -	
PS	Claim 11; SEQ ID NO 16223; 3071bp + Sequence Listing; English.	
XX		
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
CC	antmo acid sequences given in AAM82170 to AAM91921. (I) have cytosratic	
CC	activity, and can be used in gene therapy and vaccine production. (II)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting	
CC	the nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/haematopoeitic-related diseases, especially	
CC	cancers and cancer metastases of haematopoeitic-derived cells. AAK64703	
CC	to AAK87594 represent human immune/haematopoeitic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	
CC	represent sequences used in the exemplification of the present invention.	
XX		
SQ	Sequence	68 AA;
QY	Query Match	100.0%; Score 22; DB 22; Length 68;
Db	Best Local Similarity	100.0%; Pred. No. 6,5e+02;
	Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
		1 QOYO 4
		29 QOYO 32
RESULT 32		
ABP31839	ABP31839 standard; Protein; 69 AA.	
AC	ABP31839;	
XX		
DT	09-JUL-2002 (first entry)	
XX		
DE	Human nuclease-like ORF812 protein, SEQ ID NO:1624.	
XX		
XX	Human, ORF; open reading frame; ORFX; drug screening; diagnosis;	
TW	disease monitoring; cytokine; cell proliferation; cell differentiation;	
TM		

KM Immune modulation; haematopoiesis regulation; tissue growth;
 KM angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KM thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KM behaviour; cancer; proliferative disorder; neurological disorder;
 KM cardiovascular disease; immune system disorder; organ transplantation;
 KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KM hypothyroidism; cholesterol ester storage disease; infection; vulvexary;
 KM vasotrophic; antipsoriatic; antidiabetic; cytostatic; nootropic;
 KM neuroprotective; antithrombotic; anticoagulant; thrombolytic;
 KM cardiant; hypotensive; anticholesteric; antiinflammatory; immunomodulator;
 KM dermatological; analgesic; virucide; antibacterial; fungicide.
 OS Homo sapiens.
 XX
 XX MO200190366-A2.
 PN
 XX 29-NOV-2001.
 PD
 XX 24-MAY-2001; 2001WO-US17076.
 PF 24-MAY-2000; 2000US-206930P.
 (CURA-) CURAGEN CORP.
 PI Leach MD, Shinkens RA;
 XX WPI; 2002-106200/14.
 DR N-PSDB; ABN75865.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 XX
 PS Claim 10; Page 672; 2508pp; English.
 XX
 XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 CC
 XX Sequence 69 AA;
 SQ

Query Match 100.0%; Score 22; DB 23; Length 69;
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QQYQ 4
 DB 64 QQYQ 67
 RESULT 33
 ABB69881
 ID ABB69881 standard; Protein; 76 AA.
 AC ABB69881;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 36435.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 KM
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL13984.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 36435; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB57737-ABR72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 76 AA;
 SQ
 Query Match 100.0%; Score 22; DB 22; Length 76;
 Best Local Similarity 100.0%; Pred. No. 7.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QQYQ 4
 DB 10 QQYQ 13
 RESULT 34
 AAU56547
 ID AAU56547 standard; Protein; 76 AA.
 XX

AC AAU56547;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #17443.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 XX 21-APR-2000; 2000US-199047P.
 XX 02-JUN-2000; 2000US-208841P.
 XX 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59577.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 XX Example 1; SEQ ID No 17742; 1069pp; English.
 PS
 SS
 CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence. For example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 76 AA;
 QY
 Query Match 100.0%; Score 22; DB 22; Length 76;
 Best Local Similarity 100.0%; Pred. No. 7.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 16 QOYQ 19
 QY 1 QOYQ 4
 ||||
 DB 16 QOYQ 19
 RESULT 35
 AAR81957
 ID AAR81957 standard; Protein; 78 AA.
 XX

AC AAR81957;
 XX
 DT 18-MAR-1996 (first entry)
 XX
 DE ORF encoded by exon 4 of rice pyruvate orthophosphate dikinase gene.
 XX
 KW rice; promoter; exon; tissue-specific; season-specific; PPK;
 KW pyruvate orthophosphate dikinase; C4 photosynthesis; C3 plant.
 XX
 OS Oryza sativa L.
 XX
 PN JP07184657-A.
 XX
 PD 25-JUL-1995.
 XX
 PF 28-DEC-1993; 93JP-0335671.
 XX
 PR 28-DEC-1993; 93JP-0335671.
 XX
 PA (NITA-) JAPAN TAFU GURASU KK.
 XX
 DR WPI: 1995-287968/38.
 DR N-PSDB; AAQ98244.
 XX
 PT Pyruvate orthophosphate dikinase (C4 photosynthesis-related) gene
 PT and promoter in rice - shows tissue-specific and seasonal specific
 PT expression.
 XX
 PS Claim 2; Page 7; 26pp; Japanese.
 XX
 CC The sequence is the open reading frame encoded by exon 4 of rice pyruvate
 CC orthophosphate dikinase (PPDK) gene, and which encodes precursor peptide.
 CC The gene shows tissue-specific (in rice leaf and rice arista) and
 CC seasonal specific expression. The PPK gene may give a C4 photosynthesis
 CC function to a C3 plant.
 XX
 SQ Sequence 78 AA;
 QY
 Query Match 100.0%; Score 22; DB 16; Length 78;
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 30 QOYQ 33
 QY 1 QOYQ 4
 ||||
 DB 30 QOYQ 33
 RESULT 36
 AA009512
 ID AA009512 standard; Protein; 78 AA.
 XX
 AC AA009512;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 23404.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX

PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI, 2001-514838/56.
 DR N-PSDB; AAI89443.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 20, SEQ ID NO 23404; 1399pp + Sequence Listing, English.
 XX
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
 CC
 XX
 SQ Sequence 78 AA;
 SO
 Query Match 100.0%; Score 22; DB 22; Length 78;
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYQ 4
 ||||
 Db 56 QOYQ 59
 RESULT 37
 ABR35150
 ID ABR35150 standard; Protein; 78 AA.
 XX
 AC ABR35150;
 XX
 DT 08-JUL-2002 (first entry)
 XX
 KW Human ORF4123 protein, SEQ ID NO:8246.
 KW Human; ORF, open reading frame; ORF; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular diseases; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antiporatic; antidiabetic; cytostatic; neurotropic;
 KW neuroprotective; antithrombotic; anticosagulant; thrombolytic;
 KW cardiant; hypotensive; anticholel; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-US17076.
 XX
 PR 24-MAY-2000; 2000US-206690P.
 XX

PA (CURA-) CURAGEN CORP.
 XX
 XX Leach MD, Shimkets RA;
 XX
 DR WPI, 2002-106200/14.
 DR N-PSDB; ABR79176.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 XX
 PS Claim 10; Page 2306; 2508pp; English.
 XX
 CC Sequences ABR1028-ABR35561 represent 454 novel human proteins
 CC designated ORF (open reading frame) 1-454, and sequences ABR75054-
 CC ABR7587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 CC
 XX
 SQ Sequence 78 AA;
 SO
 Query Match 100.0%; Score 22; DB 23; Length 78;
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYQ 4
 ||||
 Db 21 QOYQ 24
 RESULT 38
 ABR40617
 ID ABR40617 standard; Protein; 78 AA.
 XX
 AC ABR40617;
 XX
 DT 16-MAY-2003 (first entry)
 XX
 DE Zea mays oil trait related protein sequence SEQ ID NO:54.
 XX
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW lipid-like transcription factor caleosin; ATP citrate lyase; SNPI;

KM CKC-like transcription factor; antisense inhibition; co-suppression;
 transgenic plant.
 XX
 OS Zea mays.
 XX
 PN WO2003002751-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 27-JUN-2002; 2002WO-US20152.
 XX
 PR 29-JUN-2001; 2001US-301913P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famedu OO, Harwell LT,
 PI Jones TU, Kinney AJ, Klein TW, Li C, Oliveira IC, Sakai H,
 PI Shen B, Tarczy-Hartmann MC;
 PI
 MPI: 2003-201509/19.
 N-PSDB; ACC00652.

XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, calcosin-like activity, useful for altering
 PT oil phenotypes in plants such as sunflower, coconut, soybean, wheat and
 PT rice -
 XX
 PS Claim 12; Page 128; 542pp; English.

CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, lipids-like transcription factor
 CC activity, calcosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CKC-like transcription factor activity. Also described:
 CC (1) complement (II) of (I); (2) a chimeric construct (III) comprising (I)
 CC or (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and
 CC (5) oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also
 CC be used as a hybridization probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention.

Sequence 78 AA;

Query Match 100.0%; Score 22; DB 24; Length 78;
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OQYQ 4
 ||||
 DB 47 OQYQ 50

RESULT 39
 AAM27130
 ID AAM27130 standard; Protein; 80 AA.
 XX
 AC AAM27130;
 XX
 DT 11-FEB-1998 (first entry)
 XX
 DE Equine rhinovirus 1 (ERHV1) polypeptide VP4 amino acid sequence.
 XX
 KW Equine rhinovirus 1; ERHV1; foot-and-mouth disease virus; vaccine; horse;
 KW diagnosis; antigen; polypeptide; enzyme-linked immunosorbent assay;
 KW recombinant protein; VP4.
 XX

OS Equine rhinovirus 1.
 XX
 PN WO9722701-A1.
 XX
 PD 26-JUN-1997.
 XX
 PF 18-DEC-1996; 96WO-AU00815.
 XX
 PR 18-DEC-1995; 95AU-0007201.
 XX
 PA (UTME) UNIV MELBOURNE.
 XX
 PI Crabb BS, Feng L, Studdert MJ;
 PI
 DR MPI: 1997-241692/31.
 DR N-PSDB; AAT85182.
 XX
 PT Genomic sequence of equine rhinovirus 1 - and derived proteins or
 PT virus-like particles, useful in vaccines and as diagnostic agents
 XX
 PS Claim 6; Page 34; 60pp; English.

CC The present sequence represents the amino acid sequence of the VP4
 CC fragment of the polypeptide of equine rhinovirus 1 (ERHV1). The taxonomic
 CC status of ERHV1 is unclear, as physicochemical studies have shown that
 CC the nucleic acid density and base composition of ERHV1 differs from other
 CC rhinoviruses. To this end, the nucleotide sequence encoding the
 CC polypeptide of ERHV1 was deduced. Analysis of this sequence suggests that
 CC ERHV1 is more closely related to foot-and-mouth disease virus. Individual
 CC ERHV1 proteins can be used to make vaccines to protect horses (and
 CC possibly other animals) against ERHV1. Oligonucleotide primers and probes
 CC can be used for diagnosis of ERHV1 or related viruses, while antigens of
 CC the ERHV1 polypeptide can be used to detect ERHV1-specific antibodies in
 CC the blood, particularly in enzyme-linked immunosorbent assay. They can
 CC differentiate between infected animals and those vaccinated with ERHV1
 CC vaccines (the infected animals will have antibodies reactive with
 CC non-capsid proteins but vaccinated animals will not). The coding DNA
 CC region for the present sequence represents an individual gene of the
 CC virus and can be expressed in host systems to produce recombinant
 CC proteins. Virus like particles containing the individual ERHV1 proteins
 CC e.g. VP1, can also be used as vectors for delivering therapeutic or
 CC other useful agents, including vaccinating epitopes from other pathogens
 CC or reproductive hormones.

Sequence 80 AA;

Query Match 100.0%; Score 22; DB 18; Length 80;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OQYQ 4
 ||||
 DB 28 OQYQ 31

RESULT 40
 AAY70584
 ID AAY70584 standard; Protein; 80 AA.
 XX
 AC AAY70584;
 XX
 DT 04-JUL-2000 (first entry)
 XX
 DE Salmonella Pathogenicity Island 2 (SPI2) SsaB protein.
 XX
 KW Salmonella Pathogenicity Island 2 locus; SPI2 locus; antigen carrier;
 KW attenuation; gram-negative cell; vaccine; cytotoxic; virulence; tumour;
 KW anti-arteriosclerotic; anti-Alzheimer's; bactericide; hepatotropic;
 KW antiinflammatory; microbial infection; therapeutic; Salmonella infection;
 KW Helicobacter pylori; stomach cancer; Herpes virus; Chlamydia pneumoniae;
 KW Alzheimer's disease; arteriosclerosis; viral pathogen; Hepatitis virus;
 KW cervical cancer; ssab; type III secretion apparatus.
 XX

OS Salmoneilla typhimurium.
 XX
 PN WO200014240-A2.
 XX
 PD 16-MAR-2000.
 XX
 PF 03-SEP-1999; 99WO-EP06514.
 XX
 PR 04-SEP-1998; 98EP-0116827.
 XX
 PA (CREA-) CREATOGEN BIOSCIENCES GMBH.
 XX
 PI Hensel M, Guzman CA, Medina E, Apfel H, Hueck C;
 XX
 DR WPI; 2000-256988/22.
 XX
 DR N-PSDB; AAZ51996.
 XX
 PT Attenuated gram-negative Salmoneilla cells, comprising inactivated genes
 in the SPI2 locus and useful for vaccinating against a range of
 disorders associated with microbial infections such as stomach and
 cervical cancers -

PS Claim 18, Fig 23K, 180pp; English.

XX
 CC The patent discloses attenuated gram-negative cells, especially
 CC Salmoneilla, in which at least 1 gene in the Salmoneilla pathogenicity
 CC island 2 (SPI2) locus has been inactivated resulting in attenuation/
 CC reduction of virulence compared to the wild type cell. The attenuated
 CC cells are used as carriers for presenting bacterial, viral or tumour
 CC antigens to a host and are capable of expressing the antigen in a target
 CC cell, especially a macrophage. The cells may therefore be used for the
 CC preparation of a prophylactic or therapeutic composition for the
 CC treatment of a chronic disease caused by a bacterium or virus, e.g.
 CC Salmoneilla infection or a tumour. The cells may also be used to vaccinate
 CC against a range of bacterial and viral pathogens e.g. Helicobacter pylori
 CC (associated with stomach cancer), Chlamydia pneumoniae (associated with
 CC arteriosclerosis and Alzheimer's disease), Borrelia burgdorferi,
 CC Nanobacteria (found in the chronically diseased kidneys of patients
 CC with crystalline deposits), Hepatitis virus (causative agent of
 CC Hepatitis B and C and associated with liver cancer), Human papilloma
 CC virus (HPV) (associated with cervical cancer) or Herpes virus.
 CC The present sequence is the SsaB protein, a secretion apparatus of
 CC type III secretion system, from the SPI2 locus of Salmoneilla.
 CC Inactivation of the ssa gene is useful for producing the attenuated
 CC cells.

XX
 SQ Sequence 80 AA;

Query Match 100.0%; Score 22; DB 21; Length 80;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 DB 43 QOYQ 46

Search completed: August 20, 2003, 09:29:41
 Job time : 15.6082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:19:07 ; Search time 4.41237 Seconds
(without alignments)
38.357 Million cell updates/sec

Title: US-09-991-433-8

Perfect score: 22

Sequence: 1 QOYQ 4

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	13	1 US-08-209-525-40	Sequence 40, Appl
2	22	100.0	15	2 US-08-986-234-8	Sequence 8, Appl
3	22	100.0	23	3 US-08-303-569B-9	Sequence 9, Appl
4	22	100.0	46	4 US-08-807-332B-8	Sequence 8, Appl
5	22	100.0	46	5 US-09-338-876-8	Sequence 8, Appl
6	22	100.0	80	6 US-09-091-219-6	Sequence 6, Appl
7	22	100.0	80	7 US-09-660-541-6	Sequence 6, Appl
8	22	100.0	91	8 US-09-206-903A-12	Sequence 12, Appl
9	22	100.0	91	9 US-09-202-122-12	Sequence 12, Appl
10	22	100.0	91	10 US-09-206-935-22	Sequence 22, Appl
11	22	100.0	91	11 US-09-206-935-22	Sequence 22, Appl
12	22	100.0	91	12 US-09-919-622A-12	Sequence 12, Appl
13	22	100.0	96	13 US-09-328-352-6083	Sequence 6083, Ap
14	22	100.0	103	14 US-09-460-384-35	Sequence 35, Appl
15	22	100.0	107	15 US-08-561-521-6	Sequence 6, Appl
16	22	100.0	107	16 US-08-561-521-8	Sequence 34, Appl
17	22	100.0	107	17 US-08-652-558-34	Sequence 6, Appl
18	22	100.0	107	18 US-08-318-157B-6	Sequence 15, Appl
19	22	100.0	107	19 US-09-025-203-15	Sequence 15, Appl
20	22	100.0	107	20 PCT-US95-01219-6	Sequence 8, Appl
21	22	100.0	108	21 US-08-070-116A-7	Sequence 7, Appl
22	22	100.0	108	22 US-08-116-247-9	Sequence 9, Appl
23	22	100.0	108	23 US-08-602-725-29	Sequence 29, Appl
24	22	100.0	108	24 US-08-557-050-7	Sequence 5892, Ap
25	22	100.0	110	25 US-09-328-352-5892	Sequence 7851, Ap
26	22	100.0	124	26 US-09-328-352-7851	

28	22	100.0	144	4 US-09-230-078A-6	Sequence 6, Appl
29	22	100.0	165	1 US-08-233-768A-40	Sequence 40, Appl
30	22	100.0	165	4 US-09-380-882-1	Sequence 1, Appl
31	22	100.0	168	4 US-09-380-882-2	Sequence 2, Appl
32	22	100.0	168	4 US-09-380-882-3	Sequence 3, Appl
33	22	100.0	169	1 US-07-874-848B-4	Sequence 4, Appl
34	22	100.0	173	3 US-08-772-270A-10	Sequence 10, Appl
35	22	100.0	176	1 US-08-449-922-1	Sequence 1, Appl
36	22	100.0	180	4 US-09-230-078A-4	Sequence 4, Appl
37	22	100.0	192	4 US-09-397-992A-27	Sequence 27, Appl
38	22	100.0	192	4 US-09-971-843-27	Sequence 27, Appl
39	22	100.0	193	4 US-09-397-992A-26	Sequence 26, Appl
40	22	100.0	193	4 US-09-397-992A-29	Sequence 29, Appl
41	22	100.0	193	4 US-09-971-843-26	Sequence 26, Appl
42	22	100.0	193	4 US-09-971-843-29	Sequence 29, Appl
43	22	100.0	202	4 US-09-107-532A-5746	Sequence 5746, Ap
44	22	100.0	208	3 US-09-206-903A-1	Sequence 1, Appl
45	22	100.0	208	3 US-09-202-122-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-209-525-40
Sequence 40, Application US/08209525
Patent No. 5571681
GENERAL INFORMATION:
APPLICANT: Janda, Kim D.
TITLE OF INVENTION: Chemical Event Selection By Suicide
NUMBER OF INVENTIONS: Substrate Conjugates
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESS: The Scripps Research Institute
STREET: 10666 No. 5571681th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,525
FILING DATE: 10-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lewis, Donald G.
REGISTRATION NUMBER: 28,636
REFERENCE/DOCKET NUMBER: SCRF 408.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-209-525-40

Query Match 100.0%; Score 22; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
Db 1 QOYQ 4

RESULT 2
US-08-986-234-8
; Sequence 8, Application US/08986234
; Patent No. 5981706
; GENERAL INFORMATION:
; APPLICANT: Wallen, et al.
; TITLE OF INVENTION: Methods for Synthesizing Heat Shock Protein Complexes
; FILE REFERENCE: UNME-0008-1
; CURRENT APPLICATION NUMBER: US/08/986,234
; CURRENT FILING DATE: 1997-12-05
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-08-986-234-8

Query Match 100.0%; Score 22; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
DB 9 QOYO 12

RESULT 3
US-08-303-569B-9
; Sequence 9, Application US/08303569B
; Patent No. 5859205
; GENERAL INFORMATION:
; APPLICANT: Adair, John R.
; APPLICANT: Athwal, Diljeet S.
; TITLE OF INVENTION: Humanised Antibodies
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5859205rfs
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,569B
; FILING DATE: 07-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yalco
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CARP-0032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-303-569B-9

Query Match 100.0%; Score 22; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4

DB 4 QOYO 7

RESULT 4
US-08-807-332B-8
; Sequence 8, Application US/08807332B
; Patent No. 5959074
; GENERAL INFORMATION:
; APPLICANT: Dreyfus, David H.
; APPLICANT: Gelfand, Erwin W.
; TITLE OF INVENTION: PRODUCTS AND PROCESSES FOR REGULATION OF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807,332B
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kovarik, Joseph E.
; REGISTRATION NUMBER: 33,005
; REFERENCE/DOCKET NUMBER: 2879-39
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-807-332B-8

Query Match 100.0%; Score 22; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
DB 15 QOYO 18

RESULT 5
US-09-338-876-8
; Sequence 8, Application US/09338876
; Patent No. 6187584
; GENERAL INFORMATION:
; APPLICANT: Dreyfus, David H.
; APPLICANT: Gelfand, Erwin W.
; TITLE OF INVENTION: PRODUCTS AND PROCESSES FOR REGULATION OF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,876
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/807,332
FILING DATE: 28-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kovarik, Joseph E.
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2879-39
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-338-876-8

Query Match 100.0%; Score 22; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 15 QOYQ 18

RESULT 6
US-09-091-219-6
Sequence 6, Application US/09091219
Patent No. 6171592
GENERAL INFORMATION:
APPLICANT: STUDEERT, Michael J.
APPLICANT: CRABB, Brendan S.
APPLICANT: FENG, Li
TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
FILE REFERENCE: 040268/0151
CURRENT APPLICATION NUMBER: US/09/091,219
CURRENT FILING DATE: 1998-10-05
EARLIER APPLICATION NUMBER: PCT/AU96/00815
EARLIER FILING DATE: 1996-12-18
EARLIER APPLICATION NUMBER: AU PN7201
EARLIER FILING DATE: 1995-12-18
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 6
LENGTH: 80
TYPE: PRT
ORGANISM: equine rhinovirus 1
US-09-091-219-6

Query Match 100.0%; Score 22; DB 3; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 28 QOYQ 31

RESULT 7
US-09-660-541-6
Sequence 6, Application US/09660541
Patent No. 6531136

GENERAL INFORMATION:
APPLICANT: STUDEERT, Michael J.
APPLICANT: CRABB, Brendan S.
APPLICANT: FENG, Li
TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
FILE REFERENCE: 040268/0151
CURRENT APPLICATION NUMBER: US/09/660,541
CURRENT FILING DATE: 2000-09-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201
PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 6
LENGTH: 80
TYPE: PRT
ORGANISM: equine rhinovirus 1
US-09-660-541-6

Query Match 100.0%; Score 22; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 28 QOYQ 31

RESULT 8
US-09-206-903A-12
Sequence 12, Application US/09206903A
Patent No. 6200780
GENERAL INFORMATION:
APPLICANT: Chen, Jian
APPLICANT: Godowski, Paul J.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Dong-Xiao
TITLE OF INVENTION: NOVEL TYPE I INTERFERONS
FILE REFERENCE: P1224-2R1
CURRENT APPLICATION NUMBER: US/09/206,903A
CURRENT FILING DATE: 1998-12-07
PRIOR APPLICATION NUMBER: US 60/106,463
PRIOR FILING DATE: 1998-10-30
NUMBER OF SEQ ID NOS: 12
SEQ ID NO 12
LENGTH: 91
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 54
OTHER INFORMATION: unknown amino acid
US-09-206-903A-12

Query Match 100.0%; Score 22; DB 3; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 12 QOYQ 15

RESULT 9
US-09-202-122-12
Sequence 12, Application US/09202122
Patent No. 6299869
GENERAL INFORMATION:
APPLICANT: Chen, Jian
APPLICANT: Godowski, Paul
APPLICANT: Wood, William I.
APPLICANT: Zhang, Dong-Xiao

```
/ TITLE OF INVENTION: HUMAN INTERFERON-EPSILON: A TYPE I INTERFERON
/ FILE REFERENCE: P1224R2 (filed)
/ CURRENT APPLICATION NUMBER: US/09/202,122
/ CURRENT FILING DATE: 1999-03-04
/ PRIOR APPLICATION NUMBER: PCT/US98/25672
/ PRIOR FILING DATE: 1998-12-03
/ NUMBER OF SEQ ID NOS: 12
/ SEQ ID NO 12
/ LENGTH: 91
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 54
/ OTHER INFORMATION: unknown amino acid
US-09-202-122-12
```

```
Query Match          100.0%; Score 22; DB 3; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
1 QOYO 4
12 QOYO 15
```

```
RESULT 10
US-09-206-935-22
/ Sequence 22, Application US/09206935
/ Patent No. 6299877
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Jian
/ APPLICANT: Godowski, Paul
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Dong-Xiao
/ TITLE OF INVENTION: NOVEL TYPE I INTERFERONS
/ FILE REFERENCE: 11669.50US05
/ CURRENT APPLICATION NUMBER: US/09/206,935
/ CURRENT FILING DATE: 1998-12-07
/ EARLIER APPLICATION NUMBER: 60/084,045
/ EARLIER FILING DATE: 1998-05-04
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 22
/ LENGTH: 91
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-206-935-22
```

```
Query Match          100.0%; Score 22; DB 3; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QOYO 4
12 QOYO 15
```

```
RESULT 11
US-09-206-936-22
/ Sequence 22, Application US/09206936A
/ Patent No. 6300475
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Jian
/ APPLICANT: Wood, William I.
/ TITLE OF INVENTION: NO. 6300475e1 Interferon
/ FILE REFERENCE: P1224R1
/ CURRENT APPLICATION NUMBER: US/09/206,936A
/ CURRENT FILING DATE: 1998-12-07
/ EARLIER APPLICATION NUMBER: US 60/067,897
/ EARLIER FILING DATE: 1998-12-08
/ NUMBER OF SEQ ID NOS: 22
/ SEQ ID NO 22
```

```
/ LENGTH: 91
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 54
/ OTHER INFORMATION: Xaa may be any amino acid.
US-09-206-936-22
```

```
Query Match          100.0%; Score 22; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QOYO 4
12 QOYO 15
```

```
RESULT 12
US-09-919-622A-12
/ Sequence 12, Application US/09919622A
/ Patent No. 6569420
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Jian
/ APPLICANT: Godowski, Paul
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Dong-Xiao
/ TITLE OF INVENTION: HUMAN INTERFERON-EPSILON: A TYPE I INTERFERON
/ FILE REFERENCE: P1224R2C1 (replacement)
/ CURRENT APPLICATION NUMBER: US/09/919,622A
/ CURRENT FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: US 09/202122,
/ PRIOR FILING DATE: 1999-03-04
/ PRIOR APPLICATION NUMBER: PCT/US98/25672
/ PRIOR FILING DATE: 1998-12-03
/ NUMBER OF SEQ ID NOS: 12
/ SEQ ID NO 12
/ LENGTH: 91
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 54
/ OTHER INFORMATION: unknown amino acid
US-09-919-622A-12
```

```
Query Match          100.0%; Score 22; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QOYO 4
12 QOYO 15
```

```
RESULT 13
US-09-328-352-6083
/ Sequence 6083, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 6083
/ LENGTH: 96
/ TYPE: PRT
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-6083
```

Query Match 100.0%; Score 22; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||
Db 23 QOYO 26

RESULT 14
US-09-460-384-35
Sequence 35, Application US/09460384
Patent No. 6337316

GENERAL INFORMATION:
APPLICANT: EL TAYAR, Nabli
BLECHNER, Steven
JAMESON, Brad
TEPPER, Mark

TITLE OF INVENTION: CD28/CTLA-4 INHIBITING PEPTIDOMIMETICS,
PHARMACEUTICAL COMPOSITIONS THEREOF, AND METHOD OF USING
SAME

NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NETWORK, P.L.L.C.
STREET: 624 Ninth Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.

ZIP: 20001
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/460,384

FILING DATE: 13-Dec-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/12312

FILING DATE: 11-JUN-1998
APPLICATION NUMBER: US 60/049,470

FILING DATE: 12-JUN-1997
ATTORNEY/AGENT INFORMATION:

NAME: YUN, Allen C
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: EL TAYAR-1A

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197

TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids

TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-460-384-35

Query Match 100.0%; Score 22; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||
Db 88 QOYO 91

RESULT 15
US-08-561-521-6
Sequence 6, Application US/08561521
Patent No. 5840299

GENERAL INFORMATION:
APPLICANT: Bendig, Mary M.
APPLICANT: Leger, Olivier J.
APPLICANT: Saldanha, Jose

APPLICANT: Jones, S. Tarran
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte

TITLE OF INVENTION: Adhesion Molecule VLA-4
NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew

STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco

STATE: California
COUNTRY: USA

ZIP: 94105
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/561,521

FILING DATE:
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/186,269A

FILING DATE: 25-JAN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Smith, William L.
REGISTRATION NUMBER: 30,223

REFERENCE/DOCKET NUMBER: 15270-14
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-543-5043
TELEFAX: 415-543-5043

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-561-521-6

Query Match 100.0%; Score 22; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||
Db 89 QOYO 92

RESULT 16
US-08-561-521-8
Sequence 8, Application US/08561521
Patent No. 5840299

GENERAL INFORMATION:
APPLICANT: Bendig, Mary M.

APPLICANT: Leger, Olivier J.
APPLICANT: Saldanha, Jose

APPLICANT: Jones, S. Tarran
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte

TITLE OF INVENTION: Adhesion Molecule VLA-4
NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew

STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco

STATE: California
COUNTRY: USA

ZIP: 94105
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/561,521
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/186,269A
FILING DATE: 25-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William L.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 15270-14
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-561-521-8

Query Match 100.0%; Score 22; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
Db 89 QOYO 92

RESULT 17
US-08-652-558-34
Sequence 34, Application US/08652558
Patent No. 5861155
GENERAL INFORMATION:
APPLICANT: LIN, AUGUSTINE YEE-THAN
TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,558
FILING DATE: JUNE 6, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB94/00387
FILING DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKWICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-34

Query Match 100.0%; Score 22; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
Db 89 QOYO 92

RESULT 18
US-08-318-157B-6
Sequence 6, Application US/08318157B
Patent No. 5874540
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,157B
FILING DATE: 05-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18733/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-157B-6

Query Match 100.0%; Score 22; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
Db 89 QOYO 92

RESULT 19
US-09-025-203-15
Sequence 15, Application US/09025203
Patent No. 6348581
GENERAL INFORMATION:
APPLICANT: Anderson, W.H. Kerr
APPLICANT: Tempest, Philip R.

APPLICANT: Carr, Frank J.
APPLICANT: Harris, William J.
APPLICANT: Atmour, Kathryn
TITLE OF INVENTION: High Affinity Humanized Anti-TAG-72 Monoclonal Antibodies
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/025,203
CURRENT FILING DATE: 1998-02-18
EARLIER APPLICATION NUMBER: PCT US97/19641
EARLIER FILING DATE: 1997-10-30
EARLIER APPLICATION NUMBER: US 60/030,173
EARLIER FILING DATE: 1996-10-31
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Microsoft Word 97 SR-2
SEQ ID NO 15
LENGTH: 107
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: REI VL
LOCATION: 1..107
09-025-203-15

Query Match 100.0%; Score 22; DB 4; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 89 QOYQ 92

RESULT 20
PCT-US95-01219-6
Sequence 6, Application PC/TUS9501219
GENERAL INFORMATION:
APPLICANT: Bendig, Mary M.
APPLICANT: Leger, Olivier J.
APPLICANT: Saldanha, Jose
APPLICANT: Jones, S. Tarran
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
ADHESION MOLECULE VIA-4
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01219
FILING DATE: 25-JAN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/186,269
FILING DATE: 25-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William L.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 15270-14
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-01219-6

Query Match 100.0%; Score 22; DB 5; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 89 QOYQ 92

RESULT 21
PCT-US95-01219-8
Sequence 8, Application PC/TUS9501219
GENERAL INFORMATION:
APPLICANT: Bendig, Mary M.
APPLICANT: Leger, Olivier J.
APPLICANT: Saldanha, Jose
APPLICANT: Jones, S. Tarran
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
ADHESION MOLECULE VIA-4
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01219
FILING DATE: 25-JAN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/186,269
FILING DATE: 25-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William L.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 15270-14
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-01219-8

Query Match 100.0%; Score 22; DB 5; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 89 QOYQ 92

RESULT 22
US-08-070-116A-7
Sequence 7, Application US/08070116A
Patent No. 5885573
GENERAL INFORMATION:

APPLICANT: Zivlin, Robert A.
APPLICANT: Jolliffe, Linda K.
TITLE OF INVENTION: Methods and Materials For Modulation
TITLE OF INVENTION: of the Immuno-suppressive Activity and
TITLE OF INVENTION: Toxicity of Monoclonal Antibodies
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/070,116A
FILING DATE: 01-JUN-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: ARCD:082
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-070-116A-7

Query Match 100.0%; Score 22; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
Db 89 QOYQ 92

ULT 23
08-116-247-9
Sequence 9, Application US/08116247
Patent No. 5929212
GENERAL INFORMATION:
APPLICANT: Jolliffe, Linda K.
APPLICANT: Zivlin, Robert A.
APPLICANT: Adair, John R.
APPLICANT: Achwal, Diljet S.
TITLE OF INVENTION: CD3 Specific Recombinant Antibody
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSER: Woodcock Masburn Kurtz Mackiewicz & No. 5929212r18
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/116,247

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/743,377
FILING DATE: 10-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Paintin, Francis A.
REGISTRATION NUMBER: 19,386
REFERENCE/DOCKET NUMBER: CARP-0011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-116-247-9

Query Match 100.0%; Score 22; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
Db 89 QOYQ 92

RESULT 24
US-08-602-725-29
Sequence 29, Application US/08602725
Patent No. 5965710
GENERAL INFORMATION:
APPLICANT: BODMER, WALTER F
APPLICANT: DURBIN, HEIDA
APPLICANT: SNARY, DAVID
APPLICANT: STEWART, LORNA MD
APPLICANT: YOUNG, SUSAN
APPLICANT: BATES, PAUL A
TITLE OF INVENTION: MONOCLONAL ANTIBODIES FOR USE IN
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF COLORECTAL CANCER
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSER: NIXON & VANDERHAYE P. C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,725
FILING DATE: 02-FEB-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/01816
FILING DATE: 19-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9317423
FILING DATE: 21-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36663
REFERENCE/DOCKET NUMBER: 1090-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4091
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Hukan RRI light chain
US-08-602-725-29

Query Match 100.0%; Score 22; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
89 QOYQ 92

RESULT 25
US-08-557-050-7
Sequence 7, Application US/08557050
Patent No. 6491916
GENERAL INFORMATION:
APPLICANT: Bluestone, Jeffrey A.
APPLICANT: Zivni, Robert A.
APPLICANT: Jolliffe, Linda K.
TITLE OF INVENTION: METHODS FOR MODULATION OF
TITLE OF INVENTION: THE IMMUNO-SUPPRESSIVE ACTIVITY AND TOXICITY OF MONOCLONAL
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: U.S.
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,050
FILING DATE: Concurrently Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06198
FILING DATE: 01-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/070,116
FILING DATE: 01-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: ARCD:208
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-557-050-7

Query Match 100.0%; Score 22; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
89 QOYQ 92

RESULT 26
US-09-328-352-5892
Sequence 5892, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 5892
LENGTH: 110
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-5892

Query Match 100.0%; Score 22; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
91 QOYQ 94

RESULT 27
US-09-328-352-7851
Sequence 7851, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7851
LENGTH: 124
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7851

Query Match 100.0%; Score 22; DB 4; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
42 QOYQ 45

RESULT 28
US-09-230-078A-6
Sequence 6, Application US/09230078A
Patent No. 6495334
GENERAL INFORMATION:
APPLICANT: Rajasekara, Gireesh
APPLICANT: Kakambi, Nagarata V.
APPLICANT: Kapur, Vivek
TITLE OF INVENTION: RECOMBINANT SER14 FIMBRIAL PROTEIN FROM SALMONELLA
FILE REFERENCE: 600.335USWO
CURRENT APPLICATION NUMBER: US/09/230,078A
CURRENT FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: PCT/US97/12639

PRIOR FILING DATE: 1997-07-18
PRIOR APPLICATION NUMBER: 60/022,191
PRIOR FILING DATE: 1996-07-19
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 6
LENGTH: 144
TYPE: PRT
ORGANISM: Salmonella enteritidis
US-09-230-078A-6

Query Match 100.0%; Score 22; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
DB 140 00YQ 143

RESULT 29

Sequence 40, Application US/08233788A
Patent No. 5635617
GENERAL INFORMATION:
APPLICANT: Doran, James L.
APPLICANT: Kay, William W.
APPLICANT: Collins, Karen S.
APPLICANT: Cloutier, Sharon C.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTION
TITLE OF INVENTION: OF SALMONELLA
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
SERVER: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,788A
FILING DATE: 26-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: King, Joshua
REGISTRATION NUMBER: 35,570
REFERENCE/DOCKET NUMBER: 920043.403C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEDANBERRY
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-233-788A-40

Query Match 100.0%; Score 22; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
DB 161 00YQ 164

RESULT 30
US-09-380-882-1
Sequence 1, Application US/09380882
Patent No. 6572851
GENERAL INFORMATION:
APPLICANT: Muramatsu, Takashi
APPLICANT: Kadomatsu, Kenji
APPLICANT: Oda, Munehiro
APPLICANT: Ikematsu, Shinya
APPLICANT: Sakuma, Sadatoshi
TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
TITLE OF INVENTION: and Hepatopathy
FILE REFERENCE: SPO-105
CURRENT APPLICATION NUMBER: US/09/380,882
CURRENT FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent In version 3.0
SEQ ID NO 1
LENGTH: 168
TYPE: PRT
ORGANISM: Homo sapiens
US-09-380-882-1

Query Match 100.0%; Score 22; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
DB 4 00YQ 7

RESULT 31
US-09-380-882-2
Sequence 2, Application US/09380882
Patent No. 6572851
GENERAL INFORMATION:
APPLICANT: Muramatsu, Takashi
APPLICANT: Kadomatsu, Kenji
APPLICANT: Oda, Munehiro
APPLICANT: Ikematsu, Shinya
APPLICANT: Sakuma, Sadatoshi
TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
TITLE OF INVENTION: and Hepatopathy
FILE REFERENCE: SPO-105
CURRENT APPLICATION NUMBER: US/09/380,882
CURRENT FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent In version 3.0
SEQ ID NO 2
LENGTH: 168
TYPE: PRT
ORGANISM: Mus sp.
US-09-380-882-2

Query Match 100.0%; Score 22; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
DB 4 00YQ 7

RESULT 32
US-09-380-882-3
Sequence 3, Application US/09380882
Patent No. 6572851
GENERAL INFORMATION:
APPLICANT: Muramatsu, Takashi
APPLICANT: Kadomatsu, Kenji
APPLICANT: Oda, Munehiro
APPLICANT: Ikematsu, Shinya

APPLICANT: Sakuma, Sadatoshi
TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
FILE REFERENCE: SPO-105
CURRENT APPLICATION NUMBER: US/09/360,882
CURRENT FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin version 3.0
SEQ ID NO: 3
LENGTH: 168
TYPE: PRP
ORGANISM: Rattus sp.
US-09-380-882-3

Query Match
Best Local Similarity 100.0%; Score 22; DB 4; Length 168;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 00YQ 4
|||||
4 00YQ 7

RESULT 33
US-07-874-848B-4
Sequence 4, Application US/07874848B
Patent No. 5461029
GENERAL INFORMATION:
APPLICANT: Joseph Mark Backer and
APPLICANT: Michael Robert Ostrand
TITLE OF INVENTION: Method of Treating and
TITLE OF INVENTION: Preventing Viral Infections
TITLE OF INVENTION: Using HBNF and MK Protein
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Antoinette F. Kanski
ADDRESS: American Cyanamid Company
STREET: 1937 West Main Street
STREET: P.O. Box 60
CITY: Stamford
STATE: Connecticut
COUNTRY: USA
ZIP: 06904-0060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC AT
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII from IBM Displaywrite 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/874,848B
FILING DATE: 19920424
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kanski, Antoinette F.
REGISTRATION NUMBER: 34,202
REFERENCE/DOCKET NUMBER: 31850-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203 321 2455
TELEFAX: 203 321 2971
TELEX: 710 474 4059
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 Amino Acids
TYPE: AMINO ACID
TOPOLOGY: Linear
MOLECULE TYPE: Protein
US-07-874-848B-4

Query Match
Best Local Similarity 100.0%; Score 22; DB 1; Length 169;
Matches 100.0%; Pred. No. 3.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 00YQ 4
|||||
DB 4 00YQ 7

RESULT 34
US-08-772-270A-10
Sequence 10, Application US/08772270A
Patent No. 601984
GENERAL INFORMATION:
APPLICANT: MacInnes, Janet
APPLICANT: Ricciardi, Paul
APPLICANT: Mallard, Bonnie
APPLICANT: Rosendal, Soren
TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/772,270A
FILING DATE: December 23, 1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Michelle
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 6580-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Actinobacillus pleuropneumoniae
US-08-772-270A-10

Query Match
Best Local Similarity 100.0%; Score 22; DB 3; Length 173;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||||
DB 154 00YQ 157

RESULT 35
US-08-449-922-1
Sequence 1, Application US/08449922
Patent No. 5510241
GENERAL INFORMATION:
APPLICANT: THORNS, CHRISTOPHER J
TITLE OF INVENTION: METHOD OF TESTING FOR SALMONELLA
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE, P C
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR

CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,922
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08,030,208
FILING DATE: 26-MAR-1993
APPLICATION NUMBER: GB 9021290.3
FILING DATE: 01-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9022570.7
FILING DATE: 17-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9106546.6
FILING DATE: 27-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: CRAWFORD, ARTHUR R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 1498-30
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 176 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Salmonella enteritidis/Salmonella dublin
US-08-449-922-1

Query Match 100.0%; Score 22; DB 1; Length 176;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
Db 172 QOYO 175

RESULT 36
US-09-230-078A-4
Sequence 4, Application US/09230078A
Patent No. 6495334
GENERAL INFORMATION:
APPLICANT: Rajasekara, Giteesh
APPLICANT: Kakamdi, Nagarala V.
APPLICANT: Kapur, Vivek
TITLE OF INVENTION: RECOMBINANT SEF14 FIMBRIAL PROTEIN FROM SALMONELLA
FILE REFERENCE: 600.335USMO
CURRENT APPLICATION NUMBER: US/09/230,078A
CURRENT FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: PCT/US97/12639
PRIOR FILING DATE: 1997-07-18
PRIOR APPLICATION NUMBER: 60/022,191
PRIOR FILING DATE: 1996-07-19
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 180
TYPE: PRT
ORGANISM: Salmonella enteritidis
US-09-230-078A-4

Query Match 100.0%; Score 22; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
Db 176 QOYO 179

RESULT 37
US-09-397-992A-27
Sequence 27, Application US/09397992A
Patent No. 6329175
GENERAL INFORMATION:
APPLICANT: Conklin, Darrell
APPLICANT: Grant, Francis J.
APPLICANT: Rixon, Mark W.
APPLICANT: Kindsvogel, Wayne
TITLE OF INVENTION: Interferon-epsilon
FILE REFERENCE: 98-46
CURRENT APPLICATION NUMBER: US/09/397,992A
CURRENT FILING DATE: 1999-09-16
PRIOR APPLICATION NUMBER: 60/101,012
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/118,578
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: 60/142,766
PRIOR FILING DATE: 1999-07-08
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 27
LENGTH: 192
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Truncated interferon-epsilon.
US-09-397-992A-27

Query Match 100.0%; Score 22; DB 4; Length 192;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
Db 70 QOYO 73

RESULT 38
US-09-971-843-27
Sequence 27, Application US/09971843
Patent No. 6544505
GENERAL INFORMATION:
APPLICANT: Conklin, Darrell C.
APPLICANT: Grant, Francis J.
APPLICANT: Rixon, Mark W.
APPLICANT: Kindsvogel, Wayne
TITLE OF INVENTION: Interferon-epsilon
FILE REFERENCE: 98-46D1
CURRENT APPLICATION NUMBER: US/09/971,843
CURRENT FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/101,012
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/118,578
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: 60/142,766
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 09/397,992
PRIOR FILING DATE: 1999-09-16
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 27
LENGTH: 192
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Truncated interferon-epsilon.
US-09-971-843-27

Query Match 100.0%; Score 22; DB 4; Length 193;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 70 QOYQ 73

RESULT 39

US-09-397-992A-26
Sequence 26, Application US/09397992A
Patent No. 6329175

GENERAL INFORMATION:

APPLICANT: Conklin, Darrell
APPLICANT: Grant, Francis J.

APPLICANT: Rixon, Mark W.
APPLICANT: Kindsvogel, Wayne

TITLE OF INVENTION: Interferon-epsilon

FILE REFERENCE: 98-46
CURRENT APPLICATION NUMBER: US/09/397,992A

CURRENT FILING DATE: 1999-09-16
PRIOR APPLICATION NUMBER: 60/101,012

PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/118,578

PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: 60/142,766

PRIOR FILING DATE: 1999-07-08
NUMBER OF SEQ ID NOS: 33

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 26

LENGTH: 193
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Truncated interferon-epsilon.
US-09-397-992A-26

Query Match 100.0%; Score 22; DB 4; Length 193;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 70 QOYQ 73

RESULT 40

US-09-397-992A-29
Sequence 29, Application US/09397992A
Patent No. 6329175

GENERAL INFORMATION:

APPLICANT: Conklin, Darrell
APPLICANT: Grant, Francis J.

APPLICANT: Rixon, Mark W.
APPLICANT: Kindsvogel, Wayne

TITLE OF INVENTION: Interferon-epsilon

FILE REFERENCE: 98-46
CURRENT APPLICATION NUMBER: US/09/397,992A

CURRENT FILING DATE: 1999-09-16
PRIOR APPLICATION NUMBER: 60/101,012

PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/118,578

PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: 60/142,766

PRIOR FILING DATE: 1999-07-08
NUMBER OF SEQ ID NOS: 33

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 29
LENGTH: 193

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Truncated interferon-epsilon.
US-09-397-992A-29

Query Match 100.0%; Score 22; DB 4; Length 193;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 70 QOYQ 73

Search completed: August 20, 2003, 09:34:31
JOB time : 5.41237 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:30:58 ; Search time 16.2474 Seconds
(without alignments)
32.488 Million cell updates/sec

Title: US-09-991-433-8
Perfect score: 22
Sequence: 1 QOYQ 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues
Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PTC_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PTCS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	10	15	US-10-227-616-63
2	22	100.0	21	14	US-10-079-754A-11
3	22	100.0	23	11	US-09-795-515-9
4	22	100.0	32	11	US-09-832-355A-43
5	22	100.0	51	10	US-09-993-844-29
6	22	100.0	58	14	US-10-079-754A-7
7	22	100.0	58	14	US-10-079-754A-8
8	22	100.0	58	14	US-10-079-754A-10
9	22	100.0	64	9	US-09-864-761-43353
10	22	100.0	66	9	US-09-867-550-1114
11	22	100.0	78	15	US-10-180-375-38
12	22	100.0	85	9	US-09-764-860-559
13	22	100.0	85	15	US-10-074-095-559
14	22	100.0	89	12	US-10-340-578-18
15	22	100.0	91	10	US-09-919-622A-12

16	22	100.0	105	11	US-09-764-891-3873	Sequence 3873, Ap
17	22	100.0	107	9	US-09-253-794-6	Sequence 6, Appl
18	22	100.0	107	10	US-09-999-925-15	Sequence 15, Appl
19	22	100.0	107	10	US-09-999-040-15	Sequence 15, Appl
20	22	100.0	107	11	US-09-998-817-15	Sequence 15, Appl
21	22	100.0	107	11	US-09-999-021-15	Sequence 15, Appl
22	22	100.0	107	12	US-10-223-880-15	Sequence 15, Appl
23	22	100.0	107	15	US-10-040-597-15	Sequence 15, Appl
24	22	100.0	108	9	US-09-229-200A-14	Sequence 14, Appl
25	22	100.0	108	15	US-10-267-286A-7	Sequence 7, Appl
26	22	100.0	136	11	US-09-202-329-6	Sequence 6, Appl
27	22	100.0	137	9	US-09-864-761-46753	Sequence 46753, A
28	22	100.0	137	11	US-09-862-540-51	Sequence 51, Appl
29	22	100.0	141	11	US-09-202-329-7	Sequence 832, Ap
30	22	100.0	142	15	US-10-156-761-8532	Sequence 1, Appl
31	22	100.0	168	9	US-09-380-882-1	Sequence 2, Appl
32	22	100.0	168	9	US-09-380-882-2	Sequence 2, Appl
33	22	100.0	168	9	US-09-380-882-3	Sequence 3, Appl
34	22	100.0	168	9	US-09-888-035-1	Sequence 1, Appl
35	22	100.0	168	9	US-09-888-035-2	Sequence 2, Appl
36	22	100.0	168	9	US-09-888-035-3	Sequence 3, Appl
37	22	100.0	168	11	US-09-983-000A-22	Sequence 22, Appl
38	22	100.0	168	15	US-10-020-541-2	Sequence 2, Appl
39	22	100.0	168	15	US-10-171-311-194	Sequence 194, App
40	22	100.0	192	11	US-09-971-843-27	Sequence 27, Appl
41	22	100.0	193	11	US-09-971-843-26	Sequence 26, Appl
42	22	100.0	193	11	US-09-971-843-29	Sequence 29, Appl
43	22	100.0	208	10	US-09-919-622A-1	Sequence 1, Appl
44	22	100.0	208	11	US-09-971-843-2	Sequence 2, Appl
45	22	100.0	208	11	US-09-971-843-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-10-227-616-63
; Sequence 63, Application US/10227616
; Publication No. US20030099662A1
; GENERAL INFORMATION:
; APPLICANT: Boyd, Robert Simon
; APPLICANT: Stamps, Alasdair Craig
; APPLICANT: Terrett, Jonathan Alexander
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: 2543-1-028
; CURRENT APPLICATION NUMBER: US/10/227,616
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: GB 0004576.5
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: GB 0031341.1
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 63
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-227-616-63

Query Match 100.0%; Score 22; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
Db 3 QOYQ 6

RESULT 2
US-10-079-754A-11
; Sequence 11, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:

APPLICANT: Glenn, Matthew
APPLICANT: Grigor, Murray R.
APPLICANT: Molenaar, Adrian J.
APPLICANT: Davis, Stephen R.
TITLE OF INVENTION: Compositions Isolated from Bovine
FILE REFERENCE: 11000.1068
CURRENT APPLICATION NUMBER: US/10/079,754A
CURRENT FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: US 09/699,146
PRIOR FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: US 60,162,701
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: US 09/644,190
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: US 60,150,330
PRIOR FILING DATE: 1999-08-23
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 21
TYPE: PRT
ORGANISM: Bovine
US-10-079-754A-11

Query Match
Best Local Similarity 100.0%; Score 22; DB 14; Length 21;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||
DB 2 00YQ 5

RESULT 3
US-09-795-515-9
Sequence 9, Application US/09795515
Publication No. US20030039645A1
GENERAL INFORMATION:
APPLICANT: Adair, John R.
APPLICANT: Atchwal, Diljeet S.
APPLICANT: Emtage, John S.
TITLE OF INVENTION: Humanised Antibodies
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20030039645A1
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/795,515
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/846,658
FILING DATE: 01-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Tujillo, Doreen Yelko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CARP-0057
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-795-515-9

Query Match
Best Local Similarity 100.0%; Score 22; DB 11; Length 23;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||
DB 4 00YQ 7

RESULT 4
US-09-832-355A-43
Sequence 43, Application US/09832355A
Publication No. US20030027751A1
GENERAL INFORMATION:
APPLICANT: Kovesdi, Imre
APPLICANT: Keszler, Paul
TITLE OF INVENTION: VEGF FUSION PROTEINS
FILE REFERENCE: 205654
CURRENT APPLICATION NUMBER: US/09/832,355A
CURRENT FILING DATE: 2001-04-10
NUMBER OF SEQ ID NOS: 126
SOFTWARE: PatentIn version 3.0
SEQ ID NO 43
LENGTH: 32
TYPE: PRT
ORGANISM: Homo sapiens
US-09-832-355A-43

Query Match
Best Local Similarity 100.0%; Score 22; DB 11; Length 32;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||
DB 4 00YQ 7

RESULT 5
US-09-993-844-29
Sequence 29, Application US/09993844
Patent No. US20020106739A1
GENERAL INFORMATION:
APPLICANT: Oakley, Robert H.
APPLICANT: Barak, Lawrence S.
APPLICANT: Laporte, Stephanie A.
APPLICANT: Caron, Marc G.
TITLE OF INVENTION: Modified G-Protein Coupled Receptors
FILE REFERENCE: 033072-026
CURRENT APPLICATION NUMBER: US/09/993,844
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: US 60/245,772
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/260,363
PRIOR FILING DATE: 2001-01-08
NUMBER OF SEQ ID NOS: 81
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29
LENGTH: 51
TYPE: PRT
ORGANISM: Homo sapiens
US-09-993-844-29

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 51;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||

Db 32 QOYQ 35

RESULT 6

US-10-079-754A-7
; Sequence 7, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-7

Query Match 100.0%; Score 22; DB 14; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYQ 4
Db 39 QOYQ 42

RESULT 7

US-10-079-754A-8
; Sequence 8, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-8

Query Match 100.0%; Score 22; DB 14; Length 58;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYQ 4
Db 39 QOYQ 42

RESULT 8

US-10-079-754A-10
; Sequence 10, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-10

Query Match 100.0%; Score 22; DB 14; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYQ 4
Db 39 QOYQ 42

RESULT 9

US-09-864-761-43353
; Sequence 43353, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

```

PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 43353
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005410.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.95
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.76
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.89
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
OTHER INFORMATION: SWISSPROT HIT: P28691, EVALUATION 5.20e-01
US-09-864-761-43353

Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 64;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
DB 40 QOYO 43

RESULT 10
US-09-867-550-1114
Sequence 1114, Application US/09867550
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Fued.
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
FILE REFERENCE: 21402-013 (Cura-313)
CURRENT APPLICATION NUMBER: US/09/867,550
PRIORITY FILING DATE: 2001-09-20
PRIORITY APPLICATION NUMBER: USN 60/208,427
PRIORITY FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1114
LENGTH: 66
TYPE: PRT
ORGANISM: Homo sapiens

```

```

US-09-867-550-1114
Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
DB 16 QOYO 19

RESULT 11
US-10-180-375-38
Sequence 38, Application US/10180375
Publication No. US20030126638A1
GENERAL INFORMATION:
APPLICANT: Allen, William B.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Farnoud, Omolayo O.
APPLICANT: Harwell, Leslie T.
APPLICANT: Helentjaris, Timothy
APPLICANT: Li, Changjiang
APPLICANT: Lowe, Keith
APPLICANT: Oliveira, Igor Cunha
APPLICANT: Shen, Bo
APPLICANT: Tarczyński, Mitchell C.
TITLE OF INVENTION: Alteration of Oil Traits in Plants
FILE REFERENCE: BR1458 US NAI
CURRENT APPLICATION NUMBER: US/10/180,375
PRIORITY FILING DATE: 2002-06-26
NUMBER OF SEQ ID NOS: 222
SOFTWARE: Microsoft Office 97
SEQ ID NO 38
LENGTH: 78
TYPE: PRT
ORGANISM: Zea mays
US-10-180-375-38

Query Match
Best Local Similarity 100.0%; Score 22; DB 15; Length 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
DB 47 QOYO 50

RESULT 12
US-09-764-860-559
Sequence 559, Application US/09764860
Patent No. US20020094953A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC008
CURRENT APPLICATION NUMBER: US/09/764,860
PRIORITY FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1198
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 559
LENGTH: 85
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (82)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-860-559

Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 00Y0 4
|||
Db 41 00Y0 44

RESULT 13

US-10-074-095-559

Sequence 559, Application US/10074095

Publication No. US20030077704A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC008C1

CURRENT APPLICATION NUMBER: US/10/074,095

CURRENT FILING DATE: 2002-02-14

PRIOR APPLICATION NUMBER: 09/764,860

PRIOR FILING DATE: 2001-01-17

PRIOR APPLICATION NUMBER: 60/179,065

PRIOR FILING DATE: 2000-01-31

PRIOR APPLICATION NUMBER: 60/180,628

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: 60/214,886

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 60/217,487

PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/225,758

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: 60/220,963

PRIOR FILING DATE: 2000-07-26

PRIOR APPLICATION NUMBER: 60/217,496

PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/225,447

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/218,290

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/225,757

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/226,868

PRIOR FILING DATE: 2000-08-22

PRIOR APPLICATION NUMBER: 60/216,647

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: 60/225,267

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/216,880

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: 60/225,270

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/251,869

PRIOR FILING DATE: 2000-12-08

PRIOR APPLICATION NUMBER: 60/235,834

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: 60/234,274

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: 60/234,223

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: 60/228,924

PRIOR FILING DATE: 2000-08-30

PRIOR APPLICATION NUMBER: 60/224,518

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/236,369

PRIOR FILING DATE: 2000-09-29

PRIOR APPLICATION NUMBER: 60/224,519

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/220,964

PRIOR FILING DATE: 2000-07-26

PRIOR APPLICATION NUMBER: 60/241,809

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/249,299

PRIOR FILING DATE: 2000-11-17

PRIOR APPLICATION NUMBER: 60/236,327

PRIOR FILING DATE: 2000-09-29

PRIOR APPLICATION NUMBER: 60/241,785

PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: 60/225,266
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/249,218
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,208
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,213
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,212
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,207
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,245
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,244
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,217
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,211
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,215
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,264
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,214
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,297
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/232,400
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/221,242
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/232,081
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/232,080
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/231,414
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/231,244
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/233,064
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/233,063
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,397
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,399
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,401
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/241,808
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,826
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,786
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,221
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/246,475
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 60/231,243
 PRIOR FILING DATE: 2000-09-08

Query Match 100.0%; Score 22; DB 15; Length 85;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 Db 41 QOYO 44

RESULT 14

US-10-340-578-18
 ; Sequence 18, Application US/10340578
 ; Publication No. US20030153097A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
 ; APPLICANT: DESHAIES, Raymond J.
 ; APPLICANT: COPE, Gregory
 ; APPLICANT: VERMA, Rati
 ; APPLICANT: AMBROGGIO, Xavier I.
 ; TITLE OF INVENTION: MODULATION OF COP9 SIGNALSOME ISOPEPTIDASE ACTIVITY
 ; FILE REFERENCE: CITI590-1
 ; CURRENT APPLICATION NUMBER: US/10/340,578
 ; CURRENT FILING DATE: 2003-01-09
 ; PRIOR APPLICATION NUMBER: US 60/355,334
 ; PRIOR FILING DATE: 2002-02-06
 ; PRIOR APPLICATION NUMBER: US 10/047,253
 ; PRIOR FILING DATE: 2002-01-14
 ; PRIOR APPLICATION NUMBER: US 10/046,961
 ; PRIOR FILING DATE: 2002-01-14
 ; PRIOR APPLICATION NUMBER: US 60/261,314
 ; PRIOR FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: US 60/322,322
 ; PRIOR FILING DATE: 2001-09-14
 ; PRIOR APPLICATION NUMBER: US 60/322,030
 ; PRIOR FILING DATE: 2001-09-14
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 18
 ; LENGTH: 89
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (42)-(42)
 ; OTHER INFORMATION: Xaa indicates 29 amino acid residues
 US-10-340-578-18

Query Match 100.0%; Score 22; DB 12; Length 89;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 Db 74 QOYO 77

RESULT 15
 US-09-919-622A-12
 ; Sequence 12, Application US/09919622A
 ; Patent No. US20020172660A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Godowski, Paul
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Dong-Xiao
 ; TITLE OF INVENTION: HUMAN INTERFERON-BETA: A TYPE I INTERFERON
 ; FILE REFERENCE: P1224R2C1 (replacement)
 ; CURRENT APPLICATION NUMBER: US/09/919,622A
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US 09/202122,
 ; PRIOR FILING DATE: 1999-03-04
 ; PRIOR APPLICATION NUMBER: PCT/US98/25672
 ; PRIOR FILING DATE: 1998-12-03
 ; NUMBER OF SEQ ID NOS: 12
 ; SEQ ID NO 12
 ; LENGTH: 91
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 54
 ; OTHER INFORMATION: unknown amino acid
 US-09-919-622A-12

Query Match 100.0%; Score 22; DB 10; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||
DB 12 QOYO 15

RESULT 16

US-09-764-891-3873
Sequence 3873, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 3873
LENGTH: 105
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (38)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (79)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (102)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-3873

Query Match 100.0%; Score 22; DB 11; Length 105;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||
DB 13 QOYO 16

US-09-253-794-6
Sequence 6, Application US/09253794
Patent No. US20020018750A1
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
ARMOUR, Kathryn L.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
MOUSE MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/253,794
FILING DATE: 22-Feb-1999

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,157
FILING DATE: 05-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18733/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-253-794-6

Query Match 100.0%; Score 22; DB 9; Length 107;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||
DB 89 QOYO 92

RESULT 18
US-09-999-025-15
Sequence 15, Application US/09999025
Publication No. US20020183497A1
GENERAL INFORMATION:
APPLICANT: Anderson, W.H. Kerr
APPLICANT: Tempest, Philip R.
APPLICANT: Carr, Frank J.
APPLICANT: Harris, William J.
TITLE OF INVENTION: High Affinity Humanized Anti-TAG-72 Monoclonal Antibodies
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/999,025
CURRENT FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: US 09/025,203
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: PCT US97/19641
PRIOR FILING DATE: 1997-10-30
PRIOR APPLICATION NUMBER: US 60/030,173
PRIOR FILING DATE: 1996-10-31
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Microsoft Word 97 SR-2
SEQ ID NO 15
LENGTH: 107
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: REI VL
LOCATION: 1..107
US-09-999-025-15

Query Match 100.0%; Score 22; DB 10; Length 107;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||
DB 89 QOYO 92

RESULT 19
US-09-999-040-15

; Sequence 15, Application US/09999040
; Publication No. US20020193574A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, W.H. Kerr
; APPLICANT: Tempest, Philip R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Harris, William J.
; APPLICANT: Armour, Kathryn
; TITLE OF INVENTION: High Affinity Humanized Anti-TAG-72 Monoclonal Antibodies
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/999,040
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 09/999,040
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 09/025,203
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: PCT US97/19641
; PRIOR FILING DATE: 1997-10-30
; PRIOR APPLICATION NUMBER: US 60/030,173
; PRIOR FILING DATE: 1996-10-31
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Word 97 SR-2
; SEQ ID NO 15
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: REI VL
; LOCATION: 1..107
US-09-999-040-15

Query Match 100.0%; Score 22; DB 10; Length 107;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
Db 89 QOYQ 92

RESULT 20
US-09-998-817-15
; Sequence 15, Application US/09998817
; Publication No. US20030004318A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, W.H. Kerr
; APPLICANT: Tempest, Philip R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Harris, William J.
; APPLICANT: Armour, Kathryn
; TITLE OF INVENTION: High Affinity Humanized Anti-TAG-72 Monoclonal
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/998,817
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/025,203
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: PCT US97/19641
; PRIOR FILING DATE: 1997-10-30
; PRIOR APPLICATION NUMBER: US 60/030,173
; PRIOR FILING DATE: 1996-10-31
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Word 97 SR-2
; SEQ ID NO 15
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: REI VL
; LOCATION: 1..107
US-09-998-817-15

Query Match 100.0%; Score 22; DB 11; Length 107;

Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
Db 89 QOYQ 92

RESULT 21
US-09-999-021-15
; Sequence 15, Application US/09999021
; Publication No. US20030013854A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, W.H. Kerr
; APPLICANT: Tempest, Philip R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Harris, William J.
; APPLICANT: Armour, Kathryn
; TITLE OF INVENTION: High Affinity Humanized Anti-TAG-72 Monoclonal Antibodies
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/999,021
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 09/025,203
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: PCT US97/19641
; PRIOR FILING DATE: 1997-10-30
; PRIOR APPLICATION NUMBER: US 60/030,173
; PRIOR FILING DATE: 1996-10-31
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Word 97 SR-2
; SEQ ID NO 15
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: REI VL
; LOCATION: 1..107
US-09-999-021-15

Query Match 100.0%; Score 22; DB 11; Length 107;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
Db 89 QOYQ 92

RESULT 22
US-10-223-880-15
; Sequence 15, Application US/10223880
; Publication No. US20030152571A1
; GENERAL INFORMATION:
; APPLICANT: JONAK, ZDENKA
; APPLICANT: JOHANSON, KYUNG O.
; APPLICANT: TAYLOR, ALEXANDER
; TITLE OF INVENTION: ANTI-ALPHAHERA3 HUMANIZED MONOCLONAL
; FILE REFERENCE: P50629C1
; CURRENT APPLICATION NUMBER: US/10/223,880
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 09/380,910
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/04987
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/039,609
; PRIOR FILING DATE: 1997-03-12
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Represents an amino acid sequence of the modified
OTHER INFORMATION: human REI kappa chain framework
US-10-223-880-15

Query Match
Best Local Similarity 100.0%; Score 22; DB 12; Length 107;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 89 QOYQ 92

RESULT 23
US-10-040-997-15
Sequence 15, Application US/100409397
Publication No. US20030013856A1

GENERAL INFORMATION:
APPLICANT: Anderson, W.H. Kerr
APPLICANT: Tempest, Philip R.
APPLICANT: Carr, Frank J.
APPLICANT: Harris, William J.
APPLICANT: Armour, Kathryn
TITLE OF INVENTION: High Affinity Humanized Anti-TAG-72 Monoclonal Antibodies
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/040,997
CURRENT FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: US 09/025,203
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: PCT US97/19641
PRIOR FILING DATE: 1997-10-30
PRIOR APPLICATION NUMBER: US 60/030,173
PRIOR FILING DATE: 1996-10-31
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Microsoft Word 97 SR-2
SEQ ID NO 15
LENGTH: 107
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: REI VL
LOCATION: 1..107
US-10-040-997-15

Query Match
Best Local Similarity 100.0%; Score 22; DB 15; Length 107;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 89 QOYQ 92

RESULT 24
US-09-229-200A-14
Sequence 14, Application US/09229200A
Patent No. US20020099179A1

GENERAL INFORMATION:
APPLICANT: Jolliffe et al.
TITLE OF INVENTION: CD4 Specific Recombinant Antibody
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: Johnson & Johnson
STREET: One Johnson & Johnson Plaza
CITY: New Brunswick
STATE: NJ
COUNTRY: USA
ZIP: 08933-7003
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS

SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/229,200A
FILING DATE: 13-Jan-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: John W. Wallen, III
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: ORT-948
TELECOMMUNICATION INFORMATION:
TELEPHONE: (858) 784-3239
TELEFAX: (908) 524-2808
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 108
TYPE: amino acid
TOPOLOGY: linear
US-09-229-200A-14

Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 108;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 89 QOYQ 92

RESULT 25
US-10-267-286A-7
Sequence 7, Application US/10267286A
Publication No. US20030108548A1
GENERAL INFORMATION:
APPLICANT: BLUESTONE, JEFFREY A.
APPLICANT: ZIVIN, ROBERT A.
APPLICANT: JOLLIFFE, LINDA K.
TITLE OF INVENTION: METHODS AND MATERIALS FOR MODULATION OF THE
TITLE OF INVENTION: IMMUNOSUPPRESSIVE ACTIVITY AND TOXICITY OF MONOCLONAL
TITLE OF INVENTION: ANTIBODIES
FILE REFERENCE: TOLT:004USCI
CURRENT APPLICATION NUMBER: US/10/267,286A
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: 08/557,050
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: PCT/US94/06198
PRIOR FILING DATE: 1994-06-01
PRIOR APPLICATION NUMBER: 08/070,116
PRIOR FILING DATE: 1993-06-01
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 108
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-267-286A-7

Query Match
Best Local Similarity 100.0%; Score 22; DB 15; Length 108;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
|||
DB 89 QOYQ 92

RESULT 26
US-09-202-329-6
Sequence 6, Application US/09202329A
Publication No. US20030124137A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Dalton, John P
/ APPLICANT: Andrews, Stuart J
/ TITLE OF INVENTION: Vaccine containing a peroxiredoxin and/or a B-tubulin
/ FILE REFERENCE: 1181-243
/ CURRENT APPLICATION NUMBER: US/09/202,329A
/ EARLIER FILING DATE: 1999-02-19
/ EARLIER APPLICATION NUMBER: GB 9612214.8
/ EARLIER FILING DATE: 1996-06-11
/ EARLIER APPLICATION NUMBER: PCT/GB97/01573
/ EARLIER FILING DATE: 1997-06-11
/ NUMBER OF SEQ ID NOS: 21
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 136
/ TYPE: PRT
/ ORGANISM: Fasciola hepatica
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (3)
/ OTHER INFORMATION: Xaa = unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (14)
/ OTHER INFORMATION: Xaa = unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (97)
/ OTHER INFORMATION: Xaa = unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (109)
/ OTHER INFORMATION: Xaa = unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (131)
/ OTHER INFORMATION: Xaa = unknown
/ NAME/KEY: unsure
/ LOCATION: (136)
/ OTHER INFORMATION: Xaa = unknown
/ US-09-202-329-6

Query Match          100.0%; Score 22; DB 11; Length 136;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYQ 4
||||
93 QOYQ 96

RESULT 27
US-09-864-761-46753
/ Sequence 46753, Application US/09864761
/ Patent No. US20020048763a1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
/ FILE REFERENCE: Aeonica-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ PRIOR FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
```

```
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 46753
/ LENGTH: 137
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AL049834.2
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.86
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.77
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.7
/ OTHER INFORMATION: EST HUMAN HIT: BE86115.1, EVALUE 6.00e-74
/ OTHER INFORMATION: SWISSPROT HIT: P14503, EVALUE 3.60e-01
/ US-09-864-761-46753

Query Match          100.0%; Score 22; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYQ 4
||||
60 QOYQ 63

RESULT 28
US-09-862-540-51
/ Sequence 51, Application US/09862540
/ Publication No. US20030023992A1
/ GENERAL INFORMATION:
/ APPLICANT: Vogel, Gabriel
/ TITLE OF INVENTION: NO. US20030023992A1 G Protein-Coupled Receptors
/ FILE REFERENCE: 00252US1
/ CURRENT APPLICATION NUMBER: US/09/862,540
/ PRIOR FILING DATE: 2001-05-22
/ PRIOR APPLICATION NUMBER: 60/206,138
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/206,139
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/208,976
/ PRIOR FILING DATE: 2000-06-02
/ NUMBER OF SEQ ID NOS: 61
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 51
/ LENGTH: 137
```

TYPE: PRT
ORGANISM: Homo sapiens
US-09-862-540-51

Query Match
Best Local Similarity 100.0%; Score 22; DB 11; Length 137;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
1111
DB 13 QOYO 16

RESULT 29
US-09-202-329-7
Sequence 7, Application US/09202329A
Publication No. US20030124137A1
GENERAL INFORMATION:
APPLICANT: Dalton, John P
APPLICANT: Andrews, Stuart J
TITLE OF INVENTION: Vaccine containing a peroxiredoxin and/or a B-tubulin
FILE REFERENCE: 1181-243

CURRENT APPLICATION NUMBER: US/09/202,329A
CURRENT FILING DATE: 1999-02-19
EARLIER APPLICATION NUMBER: GB 9612214.8
EARLIER FILING DATE: 1996-06-11
EARLIER APPLICATION NUMBER: PCT/GB97/01573
EARLIER FILING DATE: 1997-06-11
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 7
LENGTH: 141
TYPE: PRT
ORGANISM: Fasciola hepatica
US-09-202-329-7

Query Match
Best Local Similarity 100.0%; Score 22; DB 11; Length 141;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
1111
DB 95 QOYO 98

RESULT 30
US-09-156-761-8532
Sequence 8532, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OKURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 8532
LENGTH: 142
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-8532

Query Match 100.0%; Score 22; DB 15; Length 142;

Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
1111
DB 5 QOYO 8

RESULT 31
US-09-380-882-1
Sequence 1, Application US/09380882
Patent No. US20020019333A1
GENERAL INFORMATION:
APPLICANT: Muramatsu, Takashi
APPLICANT: Muramatsu, Kenji
APPLICANT: Oda, Munehiro
APPLICANT: Ikematsu, Shinya
APPLICANT: Sakuma, Sadatoshi
TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
FILE REFERENCE: SPO-105
CURRENT APPLICATION NUMBER: US/09/380,882
CURRENT FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 168
TYPE: PRT
ORGANISM: Homo sapiens
US-09-380-882-1

Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 168;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
1111
DB 4 QOYO 7

RESULT 32
US-09-380-882-2
Sequence 2, Application US/09380882
Patent No. US20020019333A1
GENERAL INFORMATION:
APPLICANT: Muramatsu, Takashi
APPLICANT: Kadomatsu, Kenji
APPLICANT: Oda, Munehiro
APPLICANT: Ikematsu, Shinya
APPLICANT: Sakuma, Sadatoshi
TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
FILE REFERENCE: SPO-105
CURRENT APPLICATION NUMBER: US/09/380,882
CURRENT FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin version 3.0
SEQ ID NO 2
LENGTH: 168
TYPE: PRT
ORGANISM: Mus sp.
US-09-380-882-2

Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 168;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
1111
DB 4 QOYO 7

RESULT 33

```
US-09-380-882-3
; Sequence 3, Application US/09380882
; Patent No. US2002001933A1
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Takashi
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Oda, Munehiro
; APPLICANT: Ikematsu, Shinya
; APPLICANT: Sakuma, Sadatoshi
; TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
; FILE REFERENCE: SPO-105
; CURRENT APPLICATION NUMBER: US/09/380,882
; CURRENT FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Rattus sp.
; US-09-882-3

Query Match          100.0%; Score 22; DB 9; Length 168;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
   ||||
   4 QOYQ 7

Db

RESULT 34
US-09-888-035-1
; Sequence 1, Application US/09888035
; Patent No. US20020083487A1
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Takashi
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Oda, Munehiro
; APPLICANT: Ikematsu, Shinya
; APPLICANT: Sakuma, Sadatoshi
; TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
; FILE REFERENCE: SPO-105
; CURRENT APPLICATION NUMBER: US/09/888,035
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-888-035-1

Query Match          100.0%; Score 22; DB 9; Length 168;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
   ||||
   4 QOYQ 7

Db

RESULT 35
US-09-888-035-2
; Sequence 2, Application US/09888035
; Patent No. US20020083487A1
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Takashi
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Oda, Munehiro
; APPLICANT: Ikematsu, Shinya
; APPLICANT: Sakuma, Sadatoshi

; TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
; FILE REFERENCE: SPO-105
; CURRENT APPLICATION NUMBER: US/09/888,035
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Mus sp.
; US-09-888-035-2

Query Match          100.0%; Score 22; DB 9; Length 168;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
   ||||
   4 QOYQ 7

Db

RESULT 36
US-09-888-035-3
; Sequence 3, Application US/09888035
; Patent No. US20020083487A1
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Takashi
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Oda, Munehiro
; APPLICANT: Ikematsu, Shinya
; APPLICANT: Sakuma, Sadatoshi
; TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
; FILE REFERENCE: SPO-105
; CURRENT APPLICATION NUMBER: US/09/888,035
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Rattus sp.
; US-09-888-035-3

Query Match          100.0%; Score 22; DB 9; Length 168;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
   ||||
   4 QOYQ 7

Db

RESULT 37
US-09-983-000A-22
; Sequence 22, Application US/09983000A
; Publication No. US20030118585A1
; GENERAL INFORMATION:
; APPLICANT: Agcy Therapeutics
; APPLICANT: Melcher, Thorsten
; APPLICANT: Mueller, Sabine
; APPLICANT: Chin, Daniel
; TITLE OF INVENTION: USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND VISUALIZ
; FILE REFERENCE: 263/180 -- Pagleman -- AGY
; CURRENT APPLICATION NUMBER: US/09/983,000A
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 168
; TYPE: PRT
```

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(32)
OTHER INFORMATION:
NAME/KEY: GENE
LOCATION: (1)..(168)
OTHER INFORMATION: Pleiotrophin
NAME/KEY: CHAIN
LOCATION: (33)..(168)
OTHER INFORMATION: Pleiotrophin
NAME/KEY: DISULFID
LOCATION: (47)..(76)
OTHER INFORMATION: By similarity.
NAME/KEY: DISULFID
LOCATION: (55)..(85)
OTHER INFORMATION: By similarity.
NAME/KEY: DISULFID
LOCATION: (62)..(89)
OTHER INFORMATION: By similarity.
NAME/KEY: DISULFID
LOCATION: (99)..(131)
OTHER INFORMATION: By similarity.
NAME/KEY: DISULFID
LOCATION: (109)..(141)
OTHER INFORMATION: By similarity.
US-09-983-000A-22

Query Match 100.0%; Score 22; DB 11; Length 168;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 00YQ 4
|||
Db 4 00YQ 7

RESULT 38
US-10-020-541-2
Sequence 2, Application US/10020541
Publication No. US20030082183A1
GENERAL INFORMATION:
APPLICANT: Larry A. Wheeler
APPLICANT: Gerald W. Davies
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATMENT
FILE REFERENCE: 17400(BAR)
CURRENT APPLICATION NUMBER: US/10/020,541
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,850
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 168
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Homo sapiens
US-10-020-541-2

Query Match 100.0%; Score 22; DB 15; Length 168;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 00YQ 4
|||
Db 4 00YQ 7

RESULT 39
US-10-171-311-194
Sequence 194, Application US/10171311

Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumei
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Glatc, Karen
APPLICANT: Ganavavrapu, Manjula
APPLICANT: Hoersch, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171,311
PRIOR FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
NUMBER OF SEQ ID NOS: 238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 194
LENGTH: 168
TYPE: PRT
ORGANISM: Homo sapiens
US-10-171-311-194

Query Match 100.0%; Score 22; DB 15; Length 168;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 00YQ 4
|||
Db 4 00YQ 7

RESULT 40
US-09-971-843-27
Sequence 27, Application US/09971843
Publication No. US20030013162A1
GENERAL INFORMATION:
APPLICANT: Conklin, Darrell C.
APPLICANT: Grant, Francis J.
APPLICANT: Rixon, Mark W.
APPLICANT: Kindsvogel, Wayne
TITLE OF INVENTION: Interferon-epsilon
FILE REFERENCE: 98-46D1
CURRENT APPLICATION NUMBER: US/09/971,843
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/101,012
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/118,578
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: 60/142,766
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 09/397,992
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 27
LENGTH: 192
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Truncated interferon-epsilon.
US-09-971-843-27

Query Match 100.0%; Score 22; DB 11; Length 192;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Wed Aug 20 12:54:19 2003

us-09-991-433-8.rapb

Page 14

Qy 1 0000 4
|||
Db 70 0000 73

Search completed: August 20, 2003, 10:08:18
Job time: 16.2474 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 4.49485 Seconds
(without alignments)
85.581 Million cell updates/sec

Title: US-09-991-433-8

Perfect score: 22

Sequence: 1 QOYQ 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Sched: 283308 seqs, 96168682 residues

1 number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	14	2 S39932	S-allele-associate
2	22	100.0	14	2 S39930	S-allele-associate
3	22	100.0	21	2 S65611	tubulin beta chain
4	22	100.0	21	2 T04142	tubulin beta-2 cha
5	22	100.0	39	2 PQ0011	tubulin beta chain
6	22	100.0	40	2 A60645	tubulin beta chain
7	22	100.0	41	2 T14350	tubulin beta-3 cha
8	22	100.0	53	2 T03936	tubulin beta-1 cha
9	22	100.0	53	2 T03959	tubulin beta-3 cha
10	22	100.0	69	2 A11514	tubulin beta chain
11	22	100.0	77	2 S06045	tubulin beta chain
12	22	100.0	77	2 S06044	tubulin beta chain
13	22	100.0	77	2 T18001	tubulin beta chain
14	22	100.0	80	2 A18069	hypothetical prote
15	22	100.0	89	2 PQ0190	probable secretion
16	22	100.0	90	2 PQ0191	pyruvate, phosphat
17	22	100.0	98	2 P70511	probable PE protei
18	22	100.0	107	2 S52507	Ig kappa chain V-1
19	22	100.0	108	1 K1HURE	hypothetical prote
20	22	100.0	109	2 T27500	tubulin beta chain
21	22	100.0	111	2 S06011	hypothetical prote
22	22	100.0	111	2 B82830	tubulin beta chain
23	22	100.0	112	2 C72473	hypothetical prote
24	22	100.0	122	2 F84035	hypothetical prote
25	22	100.0	129	2 H83712	hypothetical prote
26	22	100.0	129	2 F83603	hypothetical prote
27	22	100.0	132	2 H83605	protein F18G5.6 (1
28	22	100.0	139	2 T39203	dnal protein - fis
29	22	100.0	140	2 F82191	hypothetical prote

30	22	100.0	141	2 T46380	hypothetical prote
31	22	100.0	143	2 A82355	hypothetical prote
32	22	100.0	146	2 T16561	hypothetical prote
33	22	100.0	149	2 AF0861	conserved hypothet
34	22	100.0	149	2 S73351	adhesin P1 homolo
35	22	100.0	159	2 H83845	stage V sporulatio
36	22	100.0	163	2 S21858	hypothetical 18.7K
37	22	100.0	165	2 A40618	fimbria, SEF14 - S
38	22	100.0	168	2 B90077	conserved hypothet
39	22	100.0	168	1 C37780	pleiotrophin precu
40	22	100.0	168	1 B37780	pleiotrophin precu
41	22	100.0	168	1 A37087	pleiotrophin precu
42	22	100.0	168	2 I53001	pleiotrophin - mou
43	22	100.0	169	2 AG2091	glucosyltransferase
44	22	100.0	171	2 B96548	probable prolins-r
45	22	100.0	173	2 A49219	RTX toxin ApxIIIA

ALIGNMENTS

RESULT 1

S39932 S-allele-associated glycoprotein S4 - Japanese pear (fragment)

N/Alternate names: ribonuclease S4

C/Species: Pyrus serotina (Japanese pear)

C/Date: 09-Jun-1994 #sequence_revision 06-Sep-1996 #text_change 13-Mar-1997

C/Accession: S39932

R/Sassa, H.; Hirano, H.; Ikehashi, H.

Mol. Gen. Genet. 241, 17-25, 1993

A/Title: Identification and characterization of stylar glycoproteins associated with sel

A/Reference number: S39930; PMID:94049676; PMID:8232200

A/Accession: S39932

A/Molecule type: protein

A/Residues: 1-14 <SAS>

C/Superfamily: Enterobacter ribonuclease

C/Keywords: glycoprotein

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 14;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2

S39930 S-allele-associated glycoprotein S5 - Japanese pear (fragment)

N/Alternate names: ribonuclease S5

C/Species: Pyrus serotina (Japanese pear)

C/Date: 23-Nov-1994 #sequence_revision 06-Sep-1996 #text_change 07-Feb-1997

C/Accession: S39930

R/Sassa, H.; Hirano, H.; Ikehashi, H.

Mol. Gen. Genet. 241, 17-25, 1993

A/Title: Identification and characterization of stylar glycoproteins associated with sel

A/Reference number: S39930; PMID:94049676; PMID:8232200

A/Accession: S39930

A/Molecule type: protein

A/Residues: 1-14 <SAS>

C/Keywords: glycoprotein

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 14;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3

S65611
tubulin beta chain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C:Accession: S65611
R:Ruediger, M.; Pleesmann, U.; Ruediger, A.H.; Weber, K.
FEBS Lett. 364, 147-151, 1995
A:Title: Beta tubulin of bull sperm is polyglycyolated.
A:Reference number: S65611, PMID:95269788; PMID:7750559
A:Accession: S65611
A:Molecule type: protein
A:Residues: 1-21 <RUE>
C:Complex: heterodimer; alpha and beta chain
C:Superfamily: tubulin
C:Keywords: GTP binding; heterodimer; microtubule

Query Match 100.0%; Score 22; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYO 4
||||
8 QOYO 11

RESULT 4

T04142
tubulin beta-2 chain - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 13-Aug-1999
C:Accession: T04142
R:Okamura, S.; Okahara, K.; Iida, T.; Tanaka, M.; Nishimura, M.
submitted to the EMBL Data Library, February 1997
A:Description: Differential oscillation of the transcript levels of beta-tubulin isoatypes
A:Reference number: Z15146
A:Accession: T04142
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-24 <OKA>
A:Cross-references: EMBL:U91563; NID:g1907396; PIDD:AMB50566.1; PID:g1907397
A:Experimental source: strain Bright Yellow 2
C:Superfamily: tubulin

Query Match 100.0%; Score 22; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYO 4
||||
2 QOYO 5

RESULT 5

PQ0011
tubulin beta chain - carrot (fragment)
C:Species: Daucus carota (carrot)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 12-Apr-1995
C:Accession: PQ0011
R:Okamura, S.; Azumano, I.
Biochem. Int. 16, 1103-1109, 1988
A:Title: Primary structure of the carboxy-terminal region of a higher plant beta-tubulin
A:Reference number: PQ0011, PMID:89025927; PMID:3178860
A:Accession: PQ0011
A:Molecule type: mRNA
A:Residues: 1-39 <OKA>
A:Experimental source: strain GD2
C:Comment: The carboxy-terminal region of tubulin is highly acidic.
C:Superfamily: tubulin

Query Match 100.0%; Score 22; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
||||
Db 15 QOYO 18

RESULT 6

A60645
tubulin beta chain - nematode (Brugia pahangi) (fragment)
C:Species: Brugia pahangi
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 17-Mar-1999
C:Accession: A60645
R:Helm, R.; Selkirk, M.E.; Bradley, J.E.; Burns, R.G.; Hamilton, A.J.; Croft, S.; Maizel
Parasite Immunol. 11, 479-502, 1989
A:Title: Localization and immunogenicity of tubulin in the filarial nematodes Brugia malayi
A:Reference number: A60645; PMID:90067628; PMID:2685715
A:Accession: A60645
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-40 <HLL>
C:Superfamily: tubulin
C:Keywords: heterodimer; microtubule

Query Match 100.0%; Score 22; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
||||
Db 15 QOYO 18

RESULT 7

T14350
tubulin beta-3 chain - carrot (fragment)
C:Species: Daucus carota (carrot)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000
C:Accession: T14350
R:Okamura, S.; Sonohara, K.; Naito, K.; Ohkawa, H.; Kuramori, S.; Tatsuta, M.; Minamison
submitted to the EMBL Data Library, July 1996
A:Description: Characterization of beta-tubulin genes of carrot.
A:Reference number: Z17999
A:Accession: T14350
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-41 <OKA>
A:Cross-references: EMBL:U64430; NID:g1553124; PID:g1553125
A:Experimental source: cultivar Kinokki
C:Comment: Tubulin is a dimer of alpha and beta chains and is found in the microtubules
changeable site on its beta chain and at a nonexchangeable site not yet identified. The
C:Complex: heterodimer; alpha and beta chain
C:Superfamily: tubulin
C:Keywords: heterodimer

Query Match 100.0%; Score 22; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
||||
Db 17 QOYO 20

RESULT 8

T03936
tubulin beta-1 chain - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 13-Aug-1999
C:Accession: T03936
R:Okamura, S.; Okahara, K.; Iida, T.; Tanaka, M.; Nishimura, M.
submitted to the EMBL Data Library, February 1997
A:Description: Differential oscillation of the transcript levels of beta-tubulin isoatypes
A:Reference number: Z15146
A:Accession: T03936

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-53 <OKA>
A:Cross-references: EMBL:U91562; NID:g1907394; PIDN:AAB50565.1; PID:g1907395
A:Experimental source: strain Bright yellow 2
C:Superfamily: tubulin

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 53;
Pred. No. 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||
Db 29 00YQ 32

RESULT 9

T03959
tubulin beta-3 chain - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 13-Aug-1999
C:Accession: T03959
R:Okamura, S.; Okamura, K.; Iida, T.; Tanaka, M.; Mishimura, M.
submitted to the EMBL Data Library, February 1997
A:Description: Differential oscillation of the transcript levels of beta-tubulin isotypes
A:Reference number: 215146
A:Accession: T03959
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-53 <OKA>
A:Cross-references: EMBL:U91564; NID:g2264322; PIDN:AAB63381.1; PID:g1907399
A:Experimental source: strain Bright Yellow 2
C:Superfamily: tubulin

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 53;
Pred. No. 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||
Db 29 00YQ 32

RESULT 10

A11514
hypothetical protein lin0657 [imported] - Listeria innocua (strain Clp11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: A11514
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karc, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kuntz, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schuler, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; PMID:21537279; PMID:11679669
A:Accession: A11514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-69 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC95889.1; PID:g16413097; GSPDB:GN00178
C:Genetics: A:Experimental source: strain Clp11262
A:Gene: lin0657

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 69;
Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||
Db 18 00YQ 21

RESULT 11

S06045
tubulin beta chain - carrot (fragment)
C:Species: Daucus carota (carrot)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 01-Dec-2000
C:Accession: S06045
R:Kaldenhoff, R.; Aleith, F.; Richter, G.
submitted to the EMBL Data Library, June 1988
A:Reference number: S06045
A:Accession: S06045
A:Molecule type: mRNA
A:Residues: 1-77 <KAL>
A:Cross-references: EMBL:X16608; EMBL:X15456; NID:g18369; PIDN:CAA34609.1; PID:g829261
C:Superfamily: tubulin

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 77;
Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||
Db 52 00YQ 55

RESULT 12

S06044
tubulin beta chain - red goosefoot (fragment)
C:Species: Chenopodium rubrum (red goosefoot)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 01-Dec-2000
C:Accession: S06044
R:Kaldenhoff, R.; Aleith, F.; Richter, G.
submitted to the EMBL Data Library, June 1988
A:Reference number: S06044
A:Accession: S06044
A:Molecule type: mRNA
A:Residues: 1-77 <KAL>
A:Cross-references: GB:X16700; EMBL:X15456; NID:g18227; PIDN:CAA34673.1; PID:g829259
C:Superfamily: tubulin

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 77;
Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
|||
Db 52 00YQ 55

RESULT 13

T18001
hypothetical protein a499L - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18001
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T18001
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <GRA>
A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96866.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics: A>Note: a499L

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 77;
Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4

Db 68 QOYQ 71

RESULT 14

AH0699

probable secretion system protein ssas [imported] - *Salmonella enterica* subsp. *enterica*
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: this species has also been called *Salmonella typhi*

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AH0699

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, R.; Davies, P.; Brown, A.; Davis, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001

A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar
A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AH0699

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-80 <PAR>

A:Cross-references: GB:S46966; NID:G257807; PIDN:AAB23732.1; PID:G257808

C:Genetics:

A:Gene: ssas

Query Match 100.0%; Score 22; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 78;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4

Db 43 QOYQ 46

RESULT 15

PQ0190

pyruvate, phosphate dikinase (EC 2.7.9.1) 1, cytosolic - maize (fragment)

C:Species: Zea mays (maize)

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 03-Jun-2002

C:Accession: PQ0190

R:Sheen, J.

Plant Cell 3, 225-245, 1991

A:Title: Molecular mechanisms underlying the differential expression of maize pyruvate,
A:Reference number: PQ0189; MUID:93005696; PMID:1668653

A:Accession: PQ0190

A:Molecule type: DNA

A:Residues: 1-89 <SHE>

A:Cross-references: GB:S46966; NID:G257807; PIDN:AAB23731.1; PID:G257808

A:Genetics:

A:Gene: cypdk2m1

A:Introns: 29/3

C:Superfamily: pyruvate,orthophosphate dikinase

C:Keywords: cytosol; transferase

Query Match 100.0%; Score 22; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 87;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4

Db 59 QOYQ 62

RESULT 16

PQ0191

pyruvate, phosphate dikinase (EC 2.7.9.1) 2, cytosolic - maize (fragment)

C:Species: Zea mays (maize)

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 03-Jun-2002

C:Accession: PQ0191

R:Sheen, J.

Plant Cell 3, 225-245, 1991

A:Title: Molecular mechanisms underlying the differential expression of maize pyruvate,

A:Reference number: PQ0189; MUID:93005696; PMID:1668653

A:Accession: PQ0191

A:Molecule type: DNA

A:Residues: 1-90 <SHE>

A:Cross-references: GB:S46967; NID:G257809; PIDN:AAB23732.1; PID:G257810

C:Genetics:

A:Gene: cypdk2m2

C:Superfamily: pyruvate,orthophosphate dikinase

C:Keywords: cytosol; transferase

Query Match 100.0%; Score 22; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 88;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4

Db 60 QOYQ 63

RESULT 17

F70511

probable PE protein - *Mycobacterium tuberculosis* (strain H37RV)

C:Species: *Mycobacterium tuberculosis*

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 03-Nov-2000

C:Accession: F70511

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holtrold, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: F70511

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-98 <COL>

A:Cross-references: GB:Z97559; GB:AL123456; NID:G3261820; PIDN:CAB10700.1; PID:E328638;

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: PE

C:Superfamily: *Mycobacterium leprae* hypothetical protein B1620_C2_218

Query Match 100.0%; Score 22; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 97;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4

Db 60 QOYQ 63

RESULT 18

S52507

probable membrane protein YDR012c - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: hypothetical protein D2880

C:Species: *Saccharomyces cerevisiae*

C>Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 19-Apr-2002

C:Accession: S52507; S67544

R:Andre, B.; Vissers, S.; Urrutiarazu, L.

submitted to the EMBL Data Library, February 1995

A:Description: The sequence of a 42 kb segment located on the left arm of chromosome IV

A:Reference number: S52492

A:Accession: S52507

A:Molecule type: DNA

A:Residues: 1-107 <AND>

A:Cross-references: EMBL:Z48432; NID:G683669; PID:G683665

A:Experimental source: strain S288C

R:Urrutiarazu, L.A.; Andre, B.; Vissers, S.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67535

A:Accession: S67544

A:Molecule type: DNA

A:Residues: 1-107 <U9R>
 A:Cross-references: EMBL:Z74060; NID:G1430974; PID:G1430975; MIPS:YDL012C
 A:Experimental source: strain S288C
 C:Genetics:
 A:Cross-references: SGD:S0002170
 A:Map position: 4L
 A:Introns: 15/3
 C:Keywords: transmembrane protein
 F:90-106/Domain: transmembrane #status predicted <TM>

Query Match 100.0%; Score 22; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 DB 50 QOYQ 53

RESULT 19

ig kappa chain V-I region (Rei) - human (tentative sequence)
 C:Species: Homo sapiens (man)
 C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Mar-2000
 C:Accession: A91663; A01873
 R:Palin, W.; Hilschmann, N.
 Hoppe-Seyler's Z. Physiol. Chem. 356, 167-191, 1975
 A>Title: Die Primärstruktur einer kristallinen monoklonalen Immunglobulin-L-Kette vom K
 vollstendige Aminosäuresequenz des Proteins.
 A:Reference number: A91663; MUID:76023758; PMID:809329
 A:Accession: A91663
 A:Molecule type: protein
 A:Residues: 1-108 <PAL>
 A:Note: the C region of this chain has the Inv (1,2) marker

R:Opp, O.; Lattman, E.B.; Schiffer, M.; Huber, R.; Palm, W.
 Biochemistry 14, 4943-4952, 1975
 A>Title: The molecular structure of a dimer composed of the variable portions of the Ben
 A:Reference number: A90393; MUID:76039968; PMID:1182131
 A:Contents: annotation; X-ray crystallography, 2.0 angstroms
 C:Comment: This is a Bence Jones protein.
 C:Genetics:
 A:Gene: GDB:IGKV1
 A:Cross-references: GDB:136264
 A:Map position: 2p12-2p12
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
 main disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into la
 superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer
 F:16-90/Domain: immunoglobulin homology <IMM>
 F:23-88/Disulfide bonds: #status experimental

Query Match 100.0%; Score 22; DB 1; Length 108;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 DB 89 QOYQ 92

RESULT 20

hypothetical protein ZC15.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27500
 R:Baeham, V.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z20377
 A:Accession: T27500
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-109 <WIL>

A:Cross-references: EMBL:Z93396; NID:e1188462; PID:CA907715.1; GSPDB:GM00023; CESP:ZC15
 A:Experimental source: clone ZC15
 C:Genetics:
 A:Gene: CESP:ZC15.8
 A:Map position: 5
 A:Introns: 35/2

Query Match 100.0%; Score 22; DB 2; Length 109;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 DB 21 QOYQ 24

RESULT 21

tubulin beta chain - great pond snail (fragment)
 C:Species: Lymnaea stagnalis (great pond snail)
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 13-Aug-1999
 C:Accession: S08011
 R:Smil, A.B.; Thijssen, S.F.T.; Geraerts, W.P.M.; van Heerikhuizen, H.
 submitted to the EMBL Data Library, June 1989
 A:Reference number: S08011
 A:Accession: S08011
 A:Molecule type: mRNA
 A:Residues: 1-111 <SMI>
 A:Cross-references: EMBL:X15542; NID:99624; PID:CA93549.1; PID:99625
 C:Superfamily: tubulin

Query Match 100.0%; Score 22; DB 2; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 DB 85 QOYQ 88

RESULT 22

hypothetical protein XF0247 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: B82830
 R:anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequer
 Nature 406, 151-157, 2000
 A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: B82830
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-111 <SIM>
 A:Cross-references: GB:AE003878; GB:AE003849; NID:G9105052; PID:AA93060.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Britones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, O.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeiri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
 M.; Tanabe, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0247

Science 282, 2012-2018, 1998
 A>Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
 A'Reference number: A75000; MUID:99069613; PMID:9851916
 A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
 A'Accession: H89605
 A>Status: preliminary
 A'Molecule type: DNA
 A'Residuals: 1-132 <STO>
 A'Cross-references: GB:chr_X; PIDN:AAA01085.1; PID:gl055096; GSPDB:GN00028; CESP:PI8G5.6
 A'Gene: FI8G5.6
 A'Map position: X

Query Match 100.0%; Score 22; DB 2; Length 132;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 00YQ 4
 ||||
 119 00YQ 122

RESULT 28
 T39203
 dnaJ protein - fission yeast (*Schizosaccharomyces pombe*)
 C'Species: Schizosaccharomyces pombe
 C'Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C'Accession: T39203
 R'Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rameperger, U.; Pohl, T.
 Submitted to the EMBL Data Library, September 1999
 A'Reference number: Z21835
 A'Accession: T39203
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A'Molecule type: DNA
 A'Residuals: 1-139 <WOO>
 A'Cross-references: EMBL:AL110469; PIDN:CAB54153.1; GSPDB:GN00066; SPDB:SPAC926.05C
 A'Experimental source: strain 972h-; cosmid c926
 C'Genetics:
 A'Gene: SPDB:SPAC926.05C
 A'Map position: 1
 A'Introns: 72/1

Query Match 100.0%; Score 22; DB 2; Length 139;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 00YQ 4
 ||||
 63 00YQ 66

Db .

RESULT 29
 F82191
 hypothetical protein VC1510 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
 C'Species: *Vibrio cholerae*
 C'Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C'Accession: F82191
 R'Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A'Reference number: A82035; MUID:20406833; PMID:10952301
 A'Accession: F82191
 A>Status: preliminary
 A'Molecule type: DNA
 A'Residuals: 1-140 <HEI>
 A'Cross-references: GB:AE004229; GB:AE003852; NID:g9656006; PIDN:AAF94665.1; GSPDB:GN001
 A'Experimental source: serogroup O1; strain N16961; biotype El Tor
 C'Genetics:
 A'Gene: VC1510
 A'Map position: 1

Query Match 100.0%; Score 22; DB 2; Length 140;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 00YQ 4
 ||||
 124 00YQ 127

Db .

RESULT 30
 T46380
 hypothetical protein DKFZp434M1819.1 - human (fragment)
 C'Species: Homo sapiens (man)
 C'Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C'Accession: T46380
 R'Ottewill, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, January 2000
 A'Reference number: Z23031
 A'Accession: T46380
 A>Status: preliminary
 A'Molecule type: mRNA
 A'Residuals: 1-141 <AAA>
 A'Cross-references: EMBL:AL137536
 A'Experimental source: adult testis; clone DKFZp434M1819
 C'Genetics:
 A>Note: DKFZp434M1819.1

Query Match 100.0%; Score 22; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 00YQ 4
 ||||
 4 00YQ 7

Db .

RESULT 31
 A82355
 hypothetical protein VC0182 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
 C'Species: *Vibrio cholerae*
 C'Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C'Accession: A82355
 R'Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A'Reference number: A82035; MUID:20406833; PMID:10952301
 A'Accession: A82355
 A>Status: preliminary
 A'Molecule type: DNA
 A'Residuals: 1-143 <HEI>
 A'Cross-references: GB:AE004108; GB:AE003852; NID:g9654578; PIDN:AAF93358.1; GSPDB:GN001
 A'Experimental source: serogroup O1; strain N16961; biotype El Tor
 C'Genetics:
 A'Gene: VC0182
 A'Map position: 1

Query Match 100.0%; Score 22; DB 2; Length 143;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 00YQ 4
 ||||
 122 00YQ 125

Db .

RESULT 32
 T16561
 hypothetical protein K0582.1 - *Caenorhabditis elegans*
 C'Species: *Caenorhabditis elegans*
 C'Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T16561
R/Chisoe, S.
submitted to the EMBL Data Library, June 1995
A/Description: The sequence of C. elegans cosmid K05B2.
A/Reference number: Z18536
A/Accession: T16561
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-146 <CHI>
A/Cross-references: EMBL:U29379; NID:g868192; PID:g868193; PIDN:AAA68728.1; CESP:K05B2.1
A/Experimental source: strain Bristol N2
C/Genetics:
A/Gene: CESP:K05B2.1
A/Introns: 17/2; 94/2

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 146;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||||
66 QOYO 69

RESULT 33
AF0861
conserved hypothetical protein STY3102 [imported] - Salmonella enterica subsp. enterica
C/Species: Salmonella enterica subsp. enterica serovar Typh
A/Note: This species has also been called Salmonella typh
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AF0861
R/Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Conerton, P.; Croft, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Mole, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov.
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AF0861
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-149 <PAR>
A/Cross-references: GB:AL513382; PIDN:CAD06077.1; PID:g16504043; GSPDB:GN00176
C/Genetics:
A/Gene: STY3102
C/Superfamily: mhc protein; flavodoxin homology
C/Keywords: flavoprotein

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 149;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||||
45 QOYO 48

RESULT 34
S73351
addressin p1 homolog C09.orf149b - Mycoplasma pneumoniae (strain ATCC 29342)
C/Species: Mycoplasma pneumoniae
A/Variety: ATCC 29342
C/Date: 26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C/Accession: S73351
R/Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A/Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A/Reference number: S73327; MUID:97105885; PMID:8948633
A/Accession: S73351
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-149 <HIM>
A/Cross-references: EMBL:AE000004; GB:U00089; NID:g1673671; PIDN:AAB95673.1; PID:g1673671

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C/Genetics:
A/Genetic code: SGC3

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 149;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||||
77 QOYO 80

RESULT 35
H83845
stage V sporulation protein AC BHI568 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: H83845
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: AB3650; MUID:20512582; PMID:11058132
A/Accession: H83845
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-159 <STO>
A/Cross-references: GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA05287.1; GSPDB:GN00
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BHI568

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 159;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||||
14 QOYO 17

RESULT 36
S21858
hypothetical 18.7K protein - Acholeplasma phage MV-L1
C/Species: Acholeplasma phage MV-L1
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C/Accession: S21858
R/Jaeger, M.; Klotz, G.
submitted to the EMBL Data Library, April 1991
A/Description: Sequence analysis of the ss-DNA of MV-L1 - an Acholeplasma virus.
A/Reference number: S21857
A/Accession: S21858
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-163 <JAB>
A/Cross-references: EMBL:X58639; NID:g58655; PIDN:CAA41649.1; PID:g58657
C/Superfamily: Acholeplasma phage MV-L1 hypothetical 18.7K protein

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 163;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
|||||
158 QOYO 161

RESULT 37
A40618
fimbria, SFR14 - Salmonella enteritidis
N/Alternate names: fimbria, SFA
C/Species: Salmonella enteritidis
C/Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999

C.Accession: A40618, A25034
 R.Cloutier, S.C.; Muller, K.H.; Doran, J.L.; Collinson, S.K.; Kay, W.W.
 J. Bacteriol. 175, 2523-2533, 1993
 A.Title: Characterization of three fimbrial genes, sefABC, of *Salmonella enteritidis*.
 A.Reference number: A40618; MUID:93239677; PMID:8097515
 A.Contents: 27655-3b
 A.Accession: A40618
 A.Status: preliminary
 A.Molecule type: nucleic acid
 A.Residues: 1-165 <CLO>
 A.Cross-references: GB:111008; NID:9310645; PIDN:AAA7219.1; PID:9310645
 A.Note: sequence extracted from NCBI backbone (NCBIN:130387, NCBI:130395)
 R.Reuter, J.; Kay, W.W.; Trust, T.J.
 J. Bacteriol. 168, 221-227, 1986
 A.Title: Purification and characterization of fimbriae from *Salmonella enteritidis*.
 A.Reference number: A25034; MUID:87008384; PMID:2875990
 A.Accession: A25034
 A.Molecule type: protein
 A.Residues: 22-78, 'X', 80-82, 'XQ', 128 <FEU>
 Query Match 100.0%; Score 22; DB 2; Length 165;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOYQ 4
 ||||
 Db 161 QOYQ 164

RESULT 38
 B50077
 Conserved hypothetical protein SA2474 (Imported) - *Staphylococcus aureus* (strain N315)
 C.Date: 10-May-2001 #sequence revision 10-May-2001 #text_change 22-Oct-2001
 C.Species: *Staphylococcus aureus*
 A.Accession: B90077
 R.Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
 ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A.Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A.Reference number: A89758; MUID:21111952; PMID:11418146
 A.Accession: B90077
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-165 <KBR>
 A.Cross-references: GB:BA000018; PID:913702640; PIDN:BA843780.1; GSPDB:GN00149
 A.Experimental source: strain N315
 A.Genetics: SA2474
 Query Match 100.0%; Score 22; DB 2; Length 165;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOYQ 4
 ||||
 Db 56 QOYQ 59

RESULT 39
 C37780
 pleiotrophin precursor - human
 N.Alternate names: ECGF-LP; HB-GAM; HBGF-8; HBNF; heparin-binding brain mitogen; heparin
 nerve growth factor; OSF-1; osteoblast specific protein
 C.Species: *Homo sapiens* (man)
 C.Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 16-Jun-2000
 C.Accession: JN0474; C37780; B37087; A42336; A6050; B42333; A42331; S14217
 R.Kretschmer, P.J.; Fairhurst, J.L.; Holmes, J.D.; Popjes, M.L.; Boehlen, P.; Koveedi, I
 Biochem. Biophys. Res. Commun. 192, 420-429, 1993
 A.Title: Genomic organization of the human HBNF gene and characterization of an HBNF var
 A.Reference number: JN0474; MUID:93249406; PMID:8484754
 A.Accession: JN0474
 A.Molecule type: DNA

A.Residues: 1-168 <KRS>
 A.Cross-references: GB:S60110; NID:9300016; PIDN:AB26456.1; PID:9300017
 R.Li, Y.S.; Milner, P.G.; Chauhan, A.K.; Watson, M.A.; Hoffman, R.M.; Kodner, C.M.; Milb
 Science 250, 1690-1694, 1990
 A.Title: Cloning and expression of a developmentally regulated protein that induces mitc
 A.Reference number: A37780; MUID:91102543; PMID:2270483
 A.Accession: C37780
 A.Status: nucleic acid sequence not shown
 A.Molecule type: mRNA
 A.Residues: 1-168 <LTA>
 A.Cross-references: GB:X52946; NID:932030; PIDN:CAA7121.1; PID:932031
 R.Tezuka, K.; Takeshita, S.; Hakeda, Y.; Kumegawa, M.; Kikuno, R.; Hashimoto-Gotoh, T.
 Biochem. Biophys. Res. Commun. 173, 246-251, 1990
 A.Title: Isolation of mouse and human cDNA clones encoding a protein expressed specifica
 A.Reference number: A37087; MUID:91076859; PMID:1701634
 A.Accession: B37087
 A.Molecule type: mRNA
 A.Residues: 1-168 <TEZ>
 A.Cross-references: GB:D90226; NID:9219946; PIDN:BA14261.1; PID:9219947
 R.Milner, P.G.; Shah, D.; Velle, R.; Donis-Keller, H.; Kumar, B.V.
 Biochemistry 31, 12023-12028, 1992
 A.Title: Cloning, nucleotide sequence, and chromosome localization of the human pleiotro
 A.Reference number: A44236; MUID:93090732; PMID:1457401
 A.Accession: A44236
 A.Molecule type: DNA
 A.Residues: 1-168 <ML>
 A.Cross-references: GB:S50405
 A.Note: sequence extracted from NCBI backbone (NCBI:119887)
 R.Huber, D.; Gautschi-Sova, P.; Boehlen, P.
 Neurochem. Res. 15, 435-439, 1990
 A.Title: Amino-terminal sequences of a novel heparin-binding protein from human, bovine,
 A.Reference number: A60650; MUID:90356159; PMID:2388713
 A.Accession: A60650
 A.Molecule type: protein
 A.Residues: 33-46, 'X', 48-54, 'X', 56-58 <HUB>
 R.Wellstein, A.; Pang, W.-J.; Khatiri, A.; Lu, Y.; Swain, S.S.; Dickson, R.B.; Sasse, J.;
 J. Biol. Chem. 267, 2582-2587, 1992
 A.Title: A heparin-binding growth factor secreted from breast cancer cells homologous to
 A.Reference number: A42333; MUID:92129946; PMID:1733956
 A.Accession: B42333
 A.Molecule type: mRNA
 A.Residues: 33-56 <WE2>
 A.Note: sequence extracted from NCBI backbone (NCBI:78671)
 A.Note: authors report sequencing the entire mRNA
 A.Accession: A42333
 A.Molecule type: protein
 A.Residues: 'X', 34-42, 'V', 44-46, 'D', 48-51 <WE2>
 A.Experimental source: cancer cells
 A.Note: sequence extracted from NCBI backbone (NCBI:78668)
 R.Kretschmer, P.J.; Fairhurst, J.L.; Decker, M.M.; Chan, C.P.; Gluzman, Y.; Boehlen, P.;
 Growth Factors 5, 99-114, 1991
 A.Title: Cloning, characterization and developmental regulation of two members of a nove
 A.Reference number: 154231; MUID:92118405; PMID:1768439
 A.Accession: 154231
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-168 <RBS>
 A.Cross-references: GB:M57399; NID:9292072; PIDN:AAA3561.1; PID:9183890
 C.Genetics: GDB:PTN; NEGF1; HBNF-1
 A.Gene: GDB:PTN; NEGF1; HBNF-1
 A.Cross-references: GDB:132913; OMIM:162095
 A.Map position: 7q33-7q33
 A.Introns: 39/1; 97/1; 151/1
 C.Superfamily: pleiotrophin
 C.Keywords: angiogenesis; brain; cytokine; growth factor; heparin binding; mitogen; neur
 F.1-33/Domains: signal sequence #status predicted <SIG>
 F.33-168/Product: pleiotrophin #status predicted <MAT>
 P.47-76, 55-85, 92-89, 99-131, 109-141/Disulfide bonds: #status predicted
 Query Match 100.0%; Score 22; DB 1; Length 168;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOYO 4
 ||||
 Db 4 QOYO 7

RESULT 40

B37780

pleiotrophin precursor - rat

N:Alternate names: ECGF-LP; HB-GAM; HBGF-8; HBNF; heparin-binding brain mitogen; heparin

OSF-1; osteoblast specific protein

C/Species: Rattus norvegicus (Norway rat)

C/Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999

C/Accession: B37780; A38292; A36159; B60650

R/Li, Y.S.; Milner, P.G.; Chauhan, A.K.; Watson, M.A.; Hoffman, R.M.; Kodner, C.M.; Milt

Science 250, 1690-1694, 1990

A/Title: Cloning and expression of a developmentally regulated protein that induces mitc

A/Reference number: A37780; PMID:9102543; PMID:2270483

A/Accession: B37780

A/Molecule type: mRNA

A/Residues: 1-166 <LTA>

A/Residues: J.I. Ravula, H.

Biochem. 265, 16721-16724, 1990

A/Title: Molecular cloning of the 18-kDa growth-associated protein of developing brain.

A/Reference number: A38292; PMID:91009080; PMID:2170351

A/Accession: A38292

A/Molecule type: mRNA

A/Residues: 1-166 <MER>

A/Cross-references: GB:M5601; GB:J05657; NID:G204571; PIDN:AAA1310.1; PID:G204572

R/Kovesdi, I.; Faltrnurel, J.L.; Kretschmer, P.J.; Boehlen, P.

Biochem. Biophys. Res. Commun. 172, 850-854, 1990

A/Title: Heparin-binding neurotrophic factor (HBNF) and MK, members of a new family of h

A/Reference number: A36159; PMID:91054508; PMID:1700712

A/Accession: A36159

A/Molecule type: protein; mRNA

A/Residues: 33-134 <KOV>

A/Cross-references: EMBL:M68916

R/Huber, D.; Gautschi-Sova, P.; Boehlen, P.

Neurochem. Res. 15, 435-439, 1990

A/Title: Amino-terminal sequences of a novel heparin-binding protein from human, bovine,

A/Reference number: A60650; PMID:90356159; PMID:2388713

A/Accession: B60650

A/Molecule type: protein

A/Residues: 33-46, 'X', 48-54 <HUB>

C/Superfamily: pleiotrophin

C/Keywords: brain; cytokine; growth factor; heparin binding; mitogen

F/1-32/Domain: signal sequence #status predicted <SIG>

F/33-168/Product: pleiotrophin #status predicted <MAT>

F/47-76, 55-85, 62-89, 99-131, 109-141/Disulfide bonds: #status predicted

Key Match

Best Local Similarity

Matches

OY 1 QOYO 4

Db 4 QOYO 7

Search completed: August 20, 2003, 09:32:44
 Job time : 6.49485 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 2.30928 Seconds
(without alignments)
81.457 Million cell updates/sec

Title: US-09-991-433-8

Perfect score: 22
Sequence: 1 QOYQ 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	108	1 KY10 HUMAN	P01607 homo sapien
2	22	100.0	111	1 TBB LYMTST	P18699 lymnaea sta
3	22	100.0	149	1 YB29 MYCPN	P75346 mycoplasma
4	22	100.0	165	1 FM SALEN	P12061 salmonella
5	22	100.0	168	1 PTN HUMAN	P21246 h pleiotrop
6	22	100.0	168	1 PTN MOUSE	P20335 mus musculu
7	22	100.0	173	1 RT3C ACTPL	Q04474 actinobacil
8	22	100.0	177	1 TBB LYMTST	P02556 lytechinus
9	22	100.0	201	1 DERM MOUSE	Q94226 mus musculu
10	22	100.0	206	1 Y086 CAEEL	P34622 caenorhabdi
11	22	100.0	215	1 PYRE SCHPO	O94331 schizosacch
12	22	100.0	219	1 GSP1 YEAST	P32835 saccharomyc
13	22	100.0	221	1 RNS2 PYRPR	O40865 pyrus pyrif
14	22	100.0	222	1 RNS3 PYRPR	O80323 pyrus pyrif
15	22	100.0	222	1 YQ04 PSEAE	Q03268 pseudomonas
16	22	100.0	223	1 Y069 YEAST	Q05015 saccharomyc
17	22	100.0	226	1 RNS7 PYRPR	O80325 pyrus pyrif
18	22	100.0	227	1 RNS5 PYRPR	P93460 pyrus pyrif
19	22	100.0	228	1 RNS1 PYRPR	O80322 pyrus pyrif
20	22	100.0	228	1 RNS4 PYRPR	O40866 pyrus pyrif
21	22	100.0	229	1 RNS6 PYRPR	O80324 pyrus pyrif
22	22	100.0	229	1 VGO7 BPP22	Q01074 bacterioph
23	22	100.0	239	1 TBB3 ANEPH	P33632 anemia phyl
24	22	100.0	240	1 DABP STRAM	O99488 staphylococ
25	22	100.0	240	1 DABP STRAU	O99411 staphylococ
26	22	100.0	240	1 DABP STRAM	O99411 staphylococ
27	22	100.0	241	1 YB46 PASMU	O99411 staphylococ
28	22	100.0	251	1 PS4L DROME	O99411 staphylococ
29	22	100.0	257	1 NEF_HV2B1	P18092 human immun
30	22	100.0	257	1 NEF_HV2B1	P18092 human immun
31	22	100.0	268	1 TR33 YEAST	O99394 saccharomyc
32	22	100.0	274	1 ABCD_HUMAN	O99411 staphylococ
33	22	100.0	274	1 PYR_FEMENI	P10652 emericella

34	22	100.0	277	1 PYR_F ASPNG	P07817 aspergillus
35	22	100.0	277	1 PYR_F ASPOR	O14416 aspergillus
36	22	100.0	278	1 PYR_F ASPRU	O13410 aspergillus
37	22	100.0	278	1 PYR_F CLAPU	O99411 staphylococ
38	22	100.0	278	1 PYR_F BENCH	P09463 penicillium
39	22	100.0	280	1 PYR_F SOLCC	O99411 staphylococ
40	22	100.0	285	1 PYR_F PARBR	O99411 staphylococ
41	22	100.0	285	1 PYR_F CYACA	O22034 cyanidium c
42	22	100.0	288	1 DSR2 HUMAN	O95456 homo sapien
43	22	100.0	291	1 CFXQ ODOGI	P49826 odontella s
44	22	100.0	292	1 TBB_STRPU	P18700 strongyloce
45	22	100.0	295	1 CFXQ_CYAME	O22025 cyanidiosch

ALIGNMENTS

RESULT 1				
KY10_HUMAN	STANDARD;	PRT;	108 AA.	
ID	KY10_HUMAN	STANDARD;	PRT;	108 AA.
AC	P01607;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Ig kappa chain V-I region Rel.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=76023758; PubMed=809329;			
RA	Palm W., Hilschmann N.;			
RT	"The primary structure of a crystalline monoclonal immunoglobulin			
RT	kappa-type L-chain, subgroup I (Bence-Jones protein Rel); isolation			
RT	and characterization of the tryptic peptides; the complete amino acid			
RT	sequence of the protein; a contribution to the elucidation of the			
RT	three-dimensional structure of antibodies, in particular their			
RT	combining site.";			
RL	Hope-Seyler's Z. Physiol. Chem. 356:167-191 (1975).			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).			
RX	MEDLINE=76039968; PubMed=1182131;			
RA	Epp O., Lattman E.E., Schiffer M., Huber R., Palm W.;			
RT	"The molecular structure of a dimer composed of the variable portions			
RT	of the Bence-Jones protein REI refined at 2.0-A resolution.";			
RL	Biochemistry 14:4943-4952 (1975).			
CC	1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)			
CC	MARKER.			
CC	1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.			
DR	PIR: A91663; KIHURE.			
DR	PDB: 1RE1; 17-FEB-84.			
DR	PDB: 1AR2; 12-NOV-97.			
DR	PDB: 1BWV; 29-DEC-99.			
DR	GO: GO:0005576; C:extracellular; NAS.			
DR	GO: GO:0008823; F:antigen binding activity; NAS.			
DR	GO: GO:0006955; P:immune response; NAS.			
DR	InterPro: IPR007110; Ig-like.			
DR	InterPro: IPR003006; Ig_MHC.			
DR	InterPro: IPR003596; Ig_V.			
DR	Pfam: PF00047; Ig_1.			
DR	SMART: SM00406; IG_V_1.			
DR	PROSITE: PS00835; IG_LIKE; 1.			
KW	Immunoglobulin V region; Bence-Jones protein; 3D-structure.			
FT	DOMAIN 1-23			
FT	DOMAIN 24-34			
FT	DOMAIN 35-49			
FT	DOMAIN 50-56			
FT	DOMAIN 57-88			
FT	DOMAIN 89-97			
FT	DOMAIN 98-107			
FT	DISULFID 23-88			
FT	STRAND 4-7			

Best Local Similarity 100.0%; Pred. No. 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
DB 85 QOYO 88

FT STRAND 10 13
FT TURN 15 16
FT STRAND 19 25
FT TURN 30 31
FT TURN 33 38
FT TURN 40 41
FT STRAND 44 49
FT TURN 50 52
FT STRAND 53 54
FT TURN 56 57
FT TURN 60 61
FT STRAND 62 67
FT TURN 68 69
FT STRAND 70 75
FT HELIX 80 82
FT STRAND 85 90
FT STRAND 97 98
FT STRAND 102 106
FT NON TER 108 108
SQ SEQUENCE 108 AA; 11902 MW; 9E8143E1188BCE2A CRC64;

Query Match 100.0%; Score 22; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
DB 89 QOYO 92

RESULT 2
TBB LYMSR
ID TBB LYMSR STANDARD; PRT; 111 AA.
AC P1859;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tubulin beta chain (Fragment).
OC Lymanaea stagnalis (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymaneoidea; Lymanaeidae; Lymanaea.
OX NCBI_Taxid=5523;
RN [1]
RP SEQUENCE FROM N.A.
RA Smit A.B., Thijssen S.F.T., Geraerts W.P.M., van Heerikhuizen H.;
RL Submitted (JUN-1989) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CHAINS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15542; CAA33549.1; -
DR PIR; S08011; S08011.
DR InterPro; IPR002453; Beta.tubulin.
DR InterPro; IPR000217; Tubulin.
DR InterPro; IPR003008; Tubulin_FtsZ.
DR Pfam; PF03953; tubulin_C; 1.
DR PROSITE; PS00227; TUBULIN, PARTIAL.
DR PROSITE; PS00228; TUBULIN_B_AUTOREG, PARTIAL.
KM Microtubules; GTP-binding.
FT NON TER 1 1
SQ SEQUENCE 111 AA; 12818 MW; CDC4PAP2BFC918E CRC64;

Query Match 100.0%; Score 22; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
DB 85 QOYO 88

RESULT 3
YB29_MYCPN
ID YB29_MYCPN STANDARD; PRT; 149 AA.
AC P75346;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MPN129 (C09_orf149b).
GN MPN129 OR MP025.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_Taxid=2104;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreid R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae"; Res. 24:4420-4449 (1996).
RL Nucleic Acids Res. 24:4420-4449 (1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: SOME, TO M.PNEUMONIAE MPN090.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000004; AAB95673.1; -
DR PIR; S73351; S73351.
KM Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 39 61 POTENTIAL.
FT TRANSMEM 82 104 POTENTIAL.
FT TRANSMEM 119 141 POTENTIAL.
SQ SEQUENCE 149 AA; 16669 MW; BB82494CCE2CAED CRC64;

Query Match 100.0%; Score 22; DB 1; Length 149;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
DB 77 QOYO 80

RESULT 4
FM_SALEN
ID FM_SALEN STANDARD; PRT; 165 AA.
AC P12061;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fimbral protein precursor.
GN SEPA OR SEPI4.
OS Salmonella enteritidis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
FT NON TER 1 1
SQ SEQUENCE FROM N.A.
RP STRAIN=27655-3B;

Query Match 100.0%; Score 22; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RA RX MEDLINE=93239677; PubMed=8097515;
RA Clouthier S.C., Mueller K.-H., Doonan J.L., Collinson S.K., Kay W.W.;
RT "Characterization of three fibrial genes, sefABC, of Salmonella
RT enteritidis."
RL J. Bacteriol. 175:2523-2533 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91072589; PubMed=1701443;
RA Thorns C.J., Sojka W.G., Chaisey D.C.;
RT "Detection of a novel fibrial structure on the surface of Salmonella
RT enteritidis by using a monoclonal antibody."
RL J. Clin. Microbiol. 28:2409-2414 (1990).
RN [3]
RP SEQUENCE FROM N.A.
RA Ogumnyl A.D., Kotlarski I., Morona R., Manning P.A.;
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 22-85.
RX MEDLINE=87008384; PubMed=2875990;
RA Feutrier J., Kay W.W., Truet T.J.;
RT "Purification and characterization of fibriae from Salmonella
RT enteritidis."
RL J. Bacteriol. 168:221-227 (1986).
RN [5]
RP FUNCTION: STRUCTURAL SUBUNIT OF THE SEP14 FIMBRIAL (S. ENTERITIDIS
RN FILAMENTOUS FIMBRIAE).
CC CC -1- SUBCELLULAR LOCATION: Fimbria.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL; L11008; AAA27219.1; -.
DR EMBL; L03833; AAA71892.1; -.
DR EMBL; X69516; CAA67141.1; -.
DR PIR; A40618; A40618.
DR PDB; 1LU0; 26-JUN-02.
KW Fimbria; Signal; 3D-structure.
FT CHAIN 1 21
FT SIGNAL 1 21
FT CONFLICT 22 165 FIMBRIAL PROTEIN.
FT CONFLICT 30 30 V -> E (IN REF. 2 AND 3).
FT CONFLICT 84 85 GA -> QW (IN REF. 4).
SQ SEQUENCE 165 AA; 16477 MW; 5B33798A3F0F9091 CRC64;
Query Match 100.0%; Score 22; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 OYOY 4
Db 161 OYOY 164

```

RP SEQUENCE FROM N.A.
RX MEDLINE=91076859; PubMed=1701634;
RA Tezuka K.I., Takeshita S., Hakeda Y., Kamegawa M., Kikuno R.,
RT Hashimoto-Gotoh T.;
RT "Isolation of mouse and human cDNA clones encoding a protein
RL expressed specifically in osteoblasts and brain tissues.";
RL Biochem. Biophys. Res. Commun. 173:246-251(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91102543; PubMed=2270483;
RA Li Y.-S., Milner P.G., Chaulan A.G., Watson M.A., Hoffman R.M.,
RT Kodner C.M., Milbrandt J., Deuel T.F.;
RT "Cloning and expression of a developmentally regulated protein that
RL induces mitogenic and neurite outgrowth activity.";
RL Science 250:1690-1694(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93249406; PubMed=8484754;
RA Kretschmer P.J., Fairhurst J.L., Holmes J.D., Peoples M.L., Boehlen P.,
RT Kovacs I.;
RT "Genomic organization of the human HBNF gene and characterization of
RL an HBNF variant protein as a splice mutant.";
RL Biochem. Biophys. Res. Commun. 192:420-429(1993).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92118405; PubMed=1768439;
RA Kretschmer P.J., Fairhurst J.L., Decker M.M., Chan C.P., Gluman Y.,
RT Boehlen P., Kovacs I.;
RT "Cloning, characterization and developmental regulation of two
RL members of a novel human gene family of neurite outgrowth-promoting
RN proteins.";
RN Growth Factors 5:99-114(1991).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=93090732; PubMed=1457401;
RA Milner P.G., Shah D., Velle R., Donis-Keller H., Kumar B.V.;
RT "Cloning, nucleotide sequence, and chromosome localization of the
RL human pleiotrophin gene.";
RL Biochemistry 31:12023-12028(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=97446018; PubMed=9299545;
RA Maeda H., Tsujimura A., Yoshitoka M., Arai Y., Kuboki Y., Mukai T.,
RT Nakamura T., Tsuji H., Nakagawa M., Hashimoto-Gotoh T.;
RT "Bone mass loss due to estrogen deficiency is compensated in
RL transgenic mice overexpressing human osteoblast stimulating
RN factor-1.";
RL Biochem. Biophys. Res. Commun. 238:528-533(1997).
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT Klausner R.D., Collins F.S., Wagner L., Shellen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.S., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udell T.B., Toshlyuki S., Carninci P., Prange C.J.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hellon E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson R.M.,
RA Rodriguez A.C., Grimmow J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Snatlus D.E.,
RT Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RN human and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[8]
 RN SEQUENCE OF 33-52.
 RP MEDLINE=92129346; PubMed=1733956;
 RX Wellstein A., Fang W., Khatri A., Lu Y., Swain S.S., Dickson R.B.,
 RA Sasse J., Riegel A.T., Lipman M.E.,
 RT "A heparin-binding growth factor secreted from breast cancer cells
 homologous to a developmentally regulated cytokine."
 RL J. Biol. Chem. 267:2582-2587(1992).
 CC
 CC -1- FUNCTION: HEPARIN BINDING MITOGENIC PROTEIN. HAS NEURITE EXTENSION
 CC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: OSTEOBLAST AND BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE PLEIOTROPHIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 EMBL, D90226; BAA14261.1; -
 EMBL, X52946; CAJ37121.1; -
 EMBL, S60110; AAB26456.1; -
 EMBL, S59641; AAB26456.1; JOINED.
 EMBL, S60111; AAB26456.1; JOINED.
 EMBL, S60736; AAB26456.1; JOINED.
 EMBL, M57399; AAA35961.1; -
 EMBL, S50409; AAB24425.1; -
 EMBL, S50405; AAB24425.1; JOINED.
 EMBL, S50408; AAB24425.1; JOINED.
 EMBL, S50394; AAB24425.1; JOINED.
 EMBL, AB004306; BAA23944.1; -
 EMBL, BC005916; AAH05916.1; -
 DR PIR, JN0474; C37780.
 DR HSSP, P21741; IMKC.
 DR GeneW, HGNC:9630; PTN.
 DR MIM, 162095; -
 DR GO, GO:0005615; C:extracellular space; TAS.
 DR GO, GO:0004864; P:protein phosphatase inhibitor activity; TAS.
 DR GO, GO:0007399; P:neurogenesis; TAS.
 DR GO, GO:0008284; P:positive regulation of cell proliferation; TAS.
 DR GO, GO:0007185; P:transmembrane receptor protein tyrosine pho. .; TAS.
 DR InterPro, IPR000762; PTN_MK.
 DR Pfam, PFO1091; PTN_MK_C_1.
 DR Pfam, PFO5196; PTN_MK_N_1.
 DR PRINTS, PRO0269; PTNMDKINE.
 DR PRODOM, PD005592; PTN_MK; 1.
 DR SMART, SM00193; PTN; 1.
 DR PROSITE, PS00619; PTN_MK_1; 1.
 DR PROSITE, PS00620; PTN_MK_2; 1.
 KM Growth factor; Mitogen, Heparin-binding; Signal.
 FT SIGNAL 1 32
 FT CHAIN 33 168
 FT DISULFID 47 76
 FT DISULFID 55 85
 FT DISULFID 62 89
 FT DISULFID 99 131
 FT DISULFID 109 141
 FT CONFLICT 42 42
 FT CONFLICT 47 47
 FT CONFLICT 47 47
 SQ SEQUENCE 168 AA; 18942 MW; 0359C08B6497B579 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 168;
 Best Local Similarity 100.0%; Pted. No. 78;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 ||||
 Db 4 QOYO 7

RESULT 6

PTN_MOUSE
 ID PTN_MOUSE STANDARD; PRT; 168 AA.
 AC P20535;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Pleiotrophin precursor (PTN) (Heparin-binding growth-associated
 DE molecule) (HB-GAM) (Heparin-binding growth factor 8) (HBGF-8)
 DE (Osteoblast specific factor 1) (OSF-1) (Heparin-binding neutrophilic
 DE factor) (HBNF).
 GN PTN.
 OS Mus musculus (Mouse), and
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090, 10116;
 RN
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC SPECIES=Rat; STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=91009080; PubMed=2170351;
 RA Meremies J., Rauvala H.;
 RT "Molecular cloning of the 18-kDa growth-associated protein of
 RT developing brain."
 RL J. Biol. Chem. 265:16721-16724(1990).
 RN [2]
 RP SEQUENCE OF 46-134 FROM N.A.
 RC SPECIES=Rat; STRAIN=Sprague-Dawley;
 RX MEDLINE=91054508; PubMed=1700712;
 RA Kovacs I., Fairhurst J.L., Kretschmer P.J., Boehlen P.;
 RT "Heparin-binding neutrophilic factor (HBNF) and MK, members of a new
 RT family of homologous, developmentally regulated proteins."
 RL Biochem. Biophys. Res. Commun. 172:850-854(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse;
 RX MEDLINE=91076859; PubMed=1701634;
 RA Tezuka K.I., Takeshita S., Hakeda Y., Kumegawa M., Kikuno R.,
 RT Hashimoto-Gotoh T.;
 RT "Isolation of mouse and human cDNA clones encoding a protein
 RT expressed specifically in osteoblasts and brain tissues."
 RL Biochem. Biophys. Res. Commun. 173:246-251(1990).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=DBA/2J; TISSUE=Liver;
 RX MEDLINE=93090271; PubMed=1457042;
 RA Katch K., Takeshita S., Sato M., Ito T., Amann E.;
 RT "Genomic organization of the mouse OSF-1 gene."
 RL DNA Cell Biol. 11:735-743(1992).
 CC
 CC -1- FUNCTION: HEPARIN BINDING MITOGENIC PROTEIN. HAS NEURITE EXTENSION
 CC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: OSTEOBLAST AND BRAIN.
 CC -1- DEVELOPMENTAL STAGE: STRONGLY EXPRESSED DURING THE EARLY POSTNATAL
 CC GROWTH PHASE OF BRAIN AS COMPARED WITH EMBRYONIC OR ADULT TISSUE.
 CC -1- SIMILARITY: BELONGS TO THE PLEIOTROPHIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL, M55601; AAA41310.1; -
 EMBL, M68916; AAA41311.1; -
 EMBL, D90225; BAA14260.1; -
 EMBL, S52357; AAB24477.1; -
 EMBL, S52338; AAB24477.1; JOINED.
 EMBL, S52345; AAB24477.1; JOINED.
 EMBL, S52353; AAB24477.1; JOINED.
 PIR, A37087; A37087.
 PIR, B37780; B37780.
 HSSP, P21741; IMKC.

	VARIANT	G -> GE (IN CLONE P-BETA-1)
172	172	

FT	172	G -> GE (TN CLONE P-BETA-1)
VARIANT	172	
NON LEN	1	

FT	VARIANT	172	172	G -> GE (IN CLONE P-BETA-1) .

SQ SEQUENCE 177 AA; 20323 MW; B2F3D8FEDCC72643 CRC64;
 Query Match 100.0%; Score 22; DB 1; Length 177;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 154 QOYO 157

RESULT 9
 DERM MOUSE
 ID DERM MOUSE STANDARD; PRT; 201 AA.
 AC Q9QZG6; Q9QZG6; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-SEP-2003 (Rel. 42, Last annotation update)
 DE Dermopontin precursor (Tyrosine-rich acidic matrix protein) (TRAMP) (Early quiescence protein-1) (EQ-1).
 DPT.
 Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 SEQUENCE FROM N.A.
 RA Tzen C.-Y., Chen M.-L.;
 RT "Cloning and characterization of early quiescence-1 (EQ-1) gene";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary, and Uterus;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S., Aizawa K., Izawa M., Ishii K., Kiyosawa H., Kondo S., Yamataka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Bateman S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Garske J., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mondavets P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohenki S., Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: SEEMS TO MEDIATE ADHESION BY CELL SURFACE INTEGRIN BINDING. MAY SERVE AS A COMMUNICATION LINK BETWEEN THE DERMAL FIBROBLAST CELL SURFACE AND ITS EXTRACELLULAR MATRIX ENVIRONMENT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By similarity).
 CC -1- PTM: TYROSINE-SULFATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE DERMATOPONTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AF143374; AAD54221.1; -;
 DR EMBL; AK019890; BAB31905.1; -;

DR MGD; MGJ:1928392; Dpt.
 KW Extracellular matrix; Cell adhesion; Repeat; Sulfation;
 KM Pyridone carboxylic acid; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 201
 FT REPEAT 26 79
 FT REPEAT 80 135
 FT REPEAT 70 75
 FT REPEAT 125 130
 FT REPEAT 181 186
 FT MOD_RES 19 19
 FT DISULFID 50 77
 FT DISULFID 90 132
 FT DISULFID 106 133
 FT DISULFID 139 196
 FT DISULFID 143 189
 FT CONFLICT 64 64
 SQ SEQUENCE 201 AA; 23995 MW; DA64550025B9C98E CRC64;
 Query Match 100.0%; Score 22; DB 1; Length 201;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 28 QOYO 31

RESULT 10
 Y086 CAEBL
 ID Y086 CAEBL STANDARD; PRT; 206 AA.
 AC P34632;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein ZK1236.6 in chromosome III.
 GN ZK1236.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; OC Rhabditidae; Felodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bontfeld J., Burton J., Connell M., Copestake T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kerhaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R., Sims M., Smailson N., Smith A., Smith M., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Spratt J., Woldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans";
 RL Nature 368:32-38(1994).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; L13200; AAA28189.1; -;
 DR PIR; S44894; S44894.
 DR WormPeP; ZK1236.6; CE00383.
 KW Hypothetical protein.

```

DR InterPro: IPR004467; Or phsipo trans.
DR InterPro: IPR002375; P-/py_2p trans.
DR InterPro: IPR000836; PRTtransferase.
DR Pfam: PF00156; Pribosyltran; 1.
DR TIGRFAMs: TIGR00336; pyre; 1.
DR PROSITE: PS00103; PUR_Pyr.PR.TRANSFER; 1.
KW Pyrimidine biosynthesis; Transferase; Glycosyltransferase.
FT ACT_SITE 102 102 BY SIMILARITY
SQ SEQUENCE 215 AA; 23716 MW; 2ECCE404DEC08DF6 CMC64;

Query Match 100.0%; Score 22; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
   |||
Db 210 QOYQ 213

RESULT 12
GSP1_YEAST STANDARD; PRT; 219 AA.
ID GSP1_YEAST STANDARD; PRT; 219 AA.
AC P32835;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE GTP-binding nuclear protein GSP1/CNR1.
GN GSP1 OR CNR1 OR CST17 OR YLR293C OR L8003.19.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid:4932;
[1]
NP SEQUENCE FROM N.A.
RX MEDLINE=93204961; PubMed=8455603;
RA Belhumeur P., Lee A., Tam R., Dipaolo T., Fortin N., Clark M.W.;
" GSP1 and GSP2, genetic suppressors of the prp20-1 mutant in
Saccharomyces cerevisiae: GTP-binding proteins involved in the
maintenance of nuclear organization.";
RL Mol. Cell. Biol. 13:2152-2161(1993).
[2]
NP SEQUENCE FROM N.A.
RX MEDLINE=93327783; PubMed=7687541;
RA Kadowaki T., Goldfarb D., Spitz L.M., Tartakoff A.M., Ohno M.;
" Regulation of RNA processing and transport by a nuclear guanine
nucleotide release protein and members of the Ras superfamily.";
RL EMBO J. 12:2929-2937(1993).
[3]
NP SEQUENCE FROM N.A.
RX STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albertman K., Andre B., Ansgore W.,
Benes V., Brueckner M., Delius H., Dubois E., Duesthenoef A.,
Entian K.-D., Floeth W., Goffeau A., Hebling U., Kneumann K.,
Leus E., Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
Louis E.-J., Messenguy F., Mews H.-W., Miosga T., Moestl D.,
Mueller-Auer S., Newtwich U., Obermayer B., Piravandi E., Pohl T.M.,
Portellelle D., Purnelle B., Reckmann S., Rieger M., Rinke M., Rose M.,
Scharfe M., Scherren B., Scholler P., Schwager C., Schwarz S.,
Underwood A.P., Uristetaru L.A., Vandenbol M., Verhasselt P.,
Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
Weller H., Zimmermann F.K., Zollner A., Hani J., Hohsels J.D.;
" The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
[4]
NP SEQUENCE OF 64-73.
RX STRAIN=S288c;
RX MEDLINE=95203288; PubMed=7895733;
RA Garrelis J.I., Fitcher B., Kobayashi R., Latter G.I., Schwender B.,
Volpe T., Warner J.R., McLaughlin C.S.;
" Protein identifications for a Saccharomyces cerevisiae protein
database.";
RL Electrophoresis 15:1466-1486(1994)

```

```

CC -1- FUNCTION: GTP-BINDING PROTEIN INVOLVED IN NUCLEOCYTOPLASMIC
CC TRANSPORT. REQUIRED FOR THE IMPORT OF PROTEIN INTO THE NUCLEUS
CC AND ALSO FOR RNA EXPORT. ESSENTIAL FOR CELL VIABILITY. BY ANALOGY
CC WITH RAS, RAN MAY BE ACTIVATED WHEN GTP IS EXCHANGED FOR BOUND GDP
CC BY RCI1 AND INACTIVATED WHEN GTP IS HYDROLYZED BY RAN UPON
CC ACTIVATION BY RANAP1.
CC -1- SUBUNIT: RAN(GSP1/CNR1) AND RANBP1 FORM A COMPLEX IN YEAST.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L08690; AAA34653.1; -
CC EMBL; X71945; CAA50747.1; -
CC EMBL; U17243; AAB67339.1; -
CC PIR; S35504; S35504.
CC HSP; P17080; IRP.
CC SWISS-2DPAGE; P32835; YEAST.
CC SGG; S0004284; GSP1.
CC DR GO; GO:0005634; C:nucleus; IDA.
CC DR GO; GO:0003929; F:RAN small monomeric GTPase activity; IGI.
CC DR GO; GO:0006997; P:nuclear organization and biogenesis; IGI.
CC DR GO; GO:0006364; P:RNA processing; IMP.
CC DR InterPro; IPR002041; RAN.
CC DR InterPro; IPR001806; Ras trnsfmg.
CC DR InterPro; IPR005225; Small_GTP.
CC Pfam; PF000071; ras; 1.
CC DR PRINTS; PRO0627; GTPRANTC4.
CC DR PRINTS; PRO0449; RASTNSFRMG.
CC DR SMART; SM00176; RAN; 1.
CC DR TIGRFAMs; TIGR00231; small_GTP; 1.
CC DR PROSITE; PS01115; RAN; 1.
CC KW GTP-binding; Nuclear protein; Protein transport.
CC NP_BIND 19 26 GTP (BY SIMILARITY).
CC FT NP_BIND 67 71 GTP (BY SIMILARITY).
CC FT NP_BIND 124 127 GTP (BY SIMILARITY).
CC SQ SEQUENCE 219 AA; 24810 MW; C9D027033F8C70E1 CRC64;

```

Query Match 100.0%; Score 22; DB 1; Length 219;

Best Local Similarity 100.0%; Pred. NO. 1e+02; Mismatches 0; Gaps 0;

1 00YQ 4
197 00YQ 200

```

RESULT 13
RNS2_PYPY STANDARD; PRT; 221 AA.
AC Q40965;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease S-2 precursor (EC 3.1.27.1) (S2-RNase).
OS Pyrus pyrifolia (Japanese pear) (Pyrus serotina).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside 1; Rosales; Rosaceae; Maloideae; Pyrus.
OX NCBI_TaxID=3767;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nijiseiki; TISSUE=style;
RA Noriooka N., Ohnishi Y., Noriooka S., Ishimizu T., Nakanishi T.,
RA Sakiyama F.;
RT "Nucleotide sequences of cDNAs encoding S2- and S4-RNases from
RT Japanese pear (Pyrus pyrifolia Nakai 1).";

```

```

RL (in) Plant Gene Register PGR95-020.
RN [2]
RP SEQUENCE OF 28-204 FROM N.A.
RC STRAIN=cv. Nijiseiki;
RX MEDLINE=97044748; PubMed=8899819;
RX Noriooka N., Noriooka S., Ohnishi Y., Ishimizu T., Oneyama C.,
RA Nakanishi T., Sakiyama F.;
RA Ishimizu T., Mitsukami Y., Shinkawa T., Natsuka S., Hase S.,
RA Miyagi M., Sakiyama F., Noriooka S.;
RT "Presence of asparagine-linked N-acetylglucosamine and chitobiose in
RT Pyrus pyrifolia S-RNase associated with gametophytic self-
RT incompatibility.";
RL Eur. J. Biochem. 263:624-634 (1999).
RN [3]
RP CARBOHYDRATE-LINKAGE SITES, AND STRUCTURE OF CARBOHYDRATES.
RC STRAIN=cv. Nijiseiki; TISSUE=style;
RX MEDLINE=99398327; PubMed=10469125;
RA Ishimizu T., Mitsukami Y., Noriooka S.,
RA Miyagi M., Sakiyama F., Shinkawa T., Natsuka S., Hase S.,
RT "Pyrus pyrifolia S-RNase associated with gametophytic self-
RT incompatibility.";
RL Eur. J. Biochem. 263:624-634 (1999).
CC -1- FUNCTION: SELF-INCOMPATIBILITY (SI) IS THE INHERITED ABILITY OF
CC A FLOWERING PLANT TO PREVENT SELF-FERTILIZATION BY DISCRIMINATING
CC BETWEEN SELF AND NON-SELF POLLEN DURING POLLINATION. IN MANY
CC SPECIES, SELF-INCOMPATIBILITY IS CONTROLLED BY
CC THE SINGLE, MULTIALLELIC LOCUS S.
CC -1- CATALYTIC ACTIVITY: Two-stage endonucleolytic cleavage to
CC nucleoside 3'-phosphates and 3'-phosphooligonucleotides with
CC 2',3'-cyclic phosphate intermediates.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- PTM: N-LINKED CORE STRUCTURE AT ASN-91 AND ASN-137 CONTAINS
CC XYLOSE. AT ASN-153 CONTAINS XYLOSE AND AT ASN-195 CONTAINS XYLOSE
CC AND FUCOSE.
CC -1- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D49527; BAA08473.1; -
CC DR HSP; P08056; IBL.
CC DR InterPro; IPR001568; RNase_T2.
CC DR Pfam; PF00445; ribonuclease_T2; 1.
CC DR PROSITE; PS00530; RNASE_T2_1; 1.
CC KW Hydroxylase; Nuclease; Endonuclease; glycoprotein; signal.
CC FT CHAIN 1 221
CC FT SIGNAL 20
CC FT DISULFID 70 112
CC FT DISULFID 176 202
CC FT ACT_SITE 55 55
CC FT ACT_SITE 105 105
CC FT ACT_SITE 109 109
CC FT CARBOHYD 91 91
CC FT CARBOHYD 137 137
CC FT CARBOHYD 153 153
CC FT CARBOHYD 195 195
CC SQ SEQUENCE 221 AA; 25750 MW; BA828175306D1727 CRC64;

```

Query Match 100.0%; Score 22; DB 1; Length 221;

Best Local Similarity 100.0%; Pred. NO. 1e+02; Mismatches 0; Gaps 0;

1 00YQ 4
30 00YQ 33

RESULT 14
RNS3_PYPY

ID	RNS3_PYPHY	STANDARD;	PRT;	222 AA.
AC	080323;			
DT	16-OCT-2001 (Rel. 40,	Created)		
DT	16-OCT-2001 (Rel. 40,	Last sequence update)		
DT	15-SEP-2003 (Rel. 42,	Last annotation update)		
DE	Ribonuclease S-3 precursor (EC 3.1.27.1) (S3-RNase)			
OS	Pyrus pyrifolia (Japanese pear) (Pyrus serotina)			
OC	Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid 1; Rosales; Rosaceae; Maloideae; Pyrus.			
OX	NCBI_TaxID=3767;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	STRAIN=cv. Hosui; TISSUE=Style;			
RX	MEDLINE=9636346; PubMed=9700066;			
RA	Ishimizu T., Shinkawa T., Sakiyama F., Noriooka S.;			
RT	"Primary structural features of roseaceous S-RNases associated with			
RT	gametophytic self-incompatibility.";			
RT	Plant Mol.. Biol. 37:931-941(1998).			
RL	[2]			
CC	CARBOHYDRATE-LINKAGE SITES, AND STRUCTURE OF CARBOHYDRATES.			
CC	STRAIN=cv. Hosui; TISSUE=Style;			
CC	MEDLINE=99396327; PubMed=10469125;			
CC	Ishimizu T., Mitsuokami Y., Shinkawa T., Natsuka S., Hase S.,			
CC	Miyagi M., Sakiyama F., Noriooka S.;			
CC	"Presence of asparagine-linked N-acetylglucosamine and chitobiose in			
CC	Pyrus pyrifolia S-RNases associated with gametophytic self-			
CC	incompatibility.";			
CC	Bur. J. Biochem. 263:624-634(1999).			
CC	-1- FUNCTION: SELF-INCOMPATIBILITY (SI) IS THE INHERITED ABILITY OF			
CC	A FLOWERING PLANT TO PREVENT SELF-FERTILIZATION BY DISCRIMINATING			
CC	BETWEEN SELF AND NON-SELF POLLEN DURING POLLINATION. IN MANY			
CC	SPECIES, SELF-INCOMPATIBILITY IS CONTROLLED BY			
CC	THE SINGLE, MULTIALLELIC LOCUS S.			
CC	-1- CATALYTIC ACTIVITY: Two-stage endonucleolytic cleavage to			
CC	nucleoside 3'-phosphates and 3'-phosphooligonucleotides with			
CC	2,3'-cyclic phosphate intermediates.			
CC	-1- PFM: N-LINKED CORE STRUCTURE AT ASN-138 CONTAINS XYLOSE.			
CC	-1- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.ebi.ac.uk/announcement/			
CC	or send an email to license@ebi.ac.uk).			
CC	-----			
DR	EMBL; AB002140; BAA32413.1; .			
DR	PIR; S39930; S39930.			
DR	PDB; 1IQO; 14-JUN-03.			
DR	InterPro; IPR001568; RNase_T2.			
DR	Pfam; PF00445; ribonuclease_T2_1.			
DR	PROSITE; PS00530; RNASE_T2_1; 1.			
FW	HydroLase; Nuclease; Endonuclease; Glycoprotein; Signal; 3D-structure.			
FT	SIGNAL	1	22	
FT	CHAIN	23	222	RIBONUCLEASE S-3.
FT	ACT_SITE	55	55	BY SIMILARITY.
FT	ACT_SITE	106	106	BY SIMILARITY.
FT	ACT_SITE	110	110	BY SIMILARITY.
FT	DISULFID	70	113	BY SIMILARITY.
FT	DISULFID	177	203	BY SIMILARITY.
FT	CARBOHYD	40	40	N-LINKED (GLCNAC. . .).
FT	CARBOHYD	138	138	N-LINKED (GLCNAC. . .).
SO	SEQUENCE	222 AA;	25747 MW;	CBB3A6FBEB53D91C1 CRC64;
QY	Query Match	100.0%;	Score 22;	DB 1; Length 222;
QY	Best Local Similarity	100.0%;	Pred. No. 1.1e+02;	
QY	Matches	4;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
db	30 QQQQ 33			

ID	Y004_PSEAE	STANDARD	PRT	222 AA
AC	003268			
DT	01-NOV-1997	(Rel. 35, Created)		
DT	01-NOV-1997	(Rel. 35, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Hypothetical protein PA2604.			
DN	PA2604			
OS	Pseudomonas aeruginosa.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;			
OC	Pseudomonadaceae; Pseudomonas.			
OX	NCBI_TaxID=287;			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=93225809; PubMed=8469112;			
RA	Hayashi T., Matsumoto H., Ohnishi M., Terawaki Y.,			
RT	"Molecular analysis of a cytotoxin-converting phage, phi CTX, of			
RT	Pseudomonas aeruginosa: structure of the attP-cos-ctx region and			
RT	integration into the serine tRNA gene."			
RL	Mol. Microbiol. 7:657-667(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Gardner R.L., Goltzy L., Tolentino E., Westbrook-Adman S., Yuan Y.,			
RA	Brody L.L., Coulter S.N., Folger K.R., Kas A., Labig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,			
RA	Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RA	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).			
CC	-1- SIMILARITY: BELONGS TO THE BIL FAMILY.			
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D13407; BAA02671.1; -			
DR	EMBL; AE0046689; AAC05992.1; -			
DR	PIR; S38835; S38835.			
DR	InterPro; IPR006213; Bax_inhbr1.			
DR	InterPro; IPR006214; UPF0005.			
DR	Pfam; PF01027; UPF0005; 1.			
DR	PROSITE; PS01243; BIL; 1.			
FW	Hypothetical protein; Transmembrane; Complete proteome.			
FT	TRANSMEM 26 46	POTENTIAL.		
FT	TRANSMEM 48 68	POTENTIAL.		
FT	TRANSMEM 75 95	POTENTIAL.		
FT	TRANSMEM 107 127	POTENTIAL.		
FT	TRANSMEM 139 159	POTENTIAL.		
FT	TRANSMEM 166 186	POTENTIAL.		
FT	TRANSMEM 198 218	POTENTIAL.		
SO	SEQUENCE 222 AA; 24106 MW; 89C468B532EB9D19 CRC64;			
Query Match	100.0%;	Score 22;	DB 1;	Length 222;
Best Local Similarity	100.0%;	Pred. No. 1.1e+02;		
Matches	4; Conservative	0; Mismatches	0; Indels	0; Gaps
Oy	1 OQYO 4			
db	4 OQYO 7			

RESULT 16
Y69 YEAST
ID Y69 YEAST STANDARD; PRT; 223 AA.
AC Q05015;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 24.5 kDa protein in ERG8-UBP8 intergenic region.
GN YMR222C OR YMR959.04C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RA PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagers K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RL Nature 387:90-93(1997).
CC -1- SIMILARITY: TO THE N-TERMINAL OF S.POMBE DIHYDROFOLATE REDUCTASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z49939; CA90193.1; --
DR PIR: S57590; S57590.
DR SGD: S0004835; YMR222C.
DR InterPro: IPR005645; DUF341.
DR Pfam: PF03959; DUF341.1.
DR Hypothetical protein.
KW SEQUENCE 223 AA; 24533 MW; F7BE3BFCEDECSB4 CRC64;
SQ
Query Match 100.0%; Score 22; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 1,le+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
DB 151 QOYO 154
RESULT 17
RNS7 PYRPY
ID RNS7 PYRPY STANDARD; PRT; 226 AA.
AC O80325;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease S-7 precursor (EC 3.1.27.1) (S7-RNase).
OS Pyrus pyrifolia (Japanese pear) (Pyrus serotina).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside I; Rosales; Rosaceae; Maloideae; Pyrus.
OX NCBI_TaxID=3767;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Okusankichi; TISSUE=style;
RX MEDLINE=98363646; PubMed=9700066;
RA Iehimizu T., Shinkawa T., Sakiyama F., Norioke S.;
RA "Primary structural features of roseaceous S-RNases associated with
RT gameophytic self-incompatibility.";
RL Plant Mol. Biol. 37:931-941(1998).
RN [2]

RP CARBOHYDRATE-LINKAGE SITES, AND STRUCTURE OF CARBOHYDRATES.
RC STRAIN=cv. Hoanji; TISSUE=style;
RX MEDLINE=99398327; PubMed=10469125.
RA Iehimizu T., Mitsuami Y., Shinkawa T., Natsuka S., Hase S.,
RA Miyagi M., Sakiyama F., Norioke S.;
RA "Presence of asparagine-linked N-acetylglucosamine and chitobiose in
RT Pyrus pyrifolia S-RNase associated with gametophytic self-
incompatibility.";
RL Eur. J. Biochem. 263:624-634(1999).
CC -1- FUNCTION: SELF-INCOMPATIBILITY (SI) IS THE INHERITED ABILITY OF
CC A FLOWERING PLANT TO PREVENT SELF-FERTILIZATION BY DISCRIMINATING
CC BETWEEN SELF AND NON-SELF POLLEN DURING POLLINATION. IN MANY
CC SPECIES, SELF-INCOMPATIBILITY IS CONTROLLED BY
CC THE SINGLE, MULTIALLELIC LOCUS S.
CC -1- CATALYTIC ACTIVITY: Two-stage endonucleolytic cleavage to
CC nucleoside 3'-phosphates and 3'-phosphooligonucleotides with
CC 2',3'-cyclic phosphate intermediates.
CC -1- PTM: N-GLYCAN AT ASN-74 AND ASN-77 CONSISTS OF MONOSACCHARIDE
CC (GLCNAC) OR DISACCHARIDE (GLCNAC-GLCNAC). N-LINKED CORE STRUCTURE
CC AT ASN-126 CONTAINS XYLOSE AND FUCOSE AND AT ASN-144 CONTAINS
CC XYLOSE. N-GLYCAN AT ASN-172 CONSISTS OF DISACCHARIDE (GLCNAC-
CC GLCNAC).
CC -1- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB02143; BAA32416.1; --
DR InterPro: IPR001568; RNase_T2.
DR Pfam: PF00445; ribonuclease_T2.1.
DR PROSITE: PS00530; RNase_T2_1; 1.
KW Hydroxylase; Nuclease; Endonuclease; Glycoprotein; signal.
FT SIGNAL 1 27
FT CHAIN 28 226
FT ACT_SITE 60 60 RIBONUCLEASE S-7.
FT ACT_SITE 112 112 BY SIMILARITY.
FT ACT_SITE 116 116 BY SIMILARITY.
FT ACT_SITE 119 119 BY SIMILARITY.
FT DISULFID 183 209 BY SIMILARITY.
FT CARBOHYD 74 74 N-LINKED (GLCNAC OR GLCNAC. . .).
FT CARBOHYD 77 77 N-LINKED (GLCNAC OR GLCNAC. . .).
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .).
SQ SEQUENCE 226 AA; 25470 MW; 75062DEAA6172714 CRC64;
Query Match 100.0%; Score 22; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 1,le+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
DB 35 QOYO 38
RESULT 18
RNS5 PYRPY
ID RNS5 PYRPY STANDARD; PRT; 227 AA.
AC P93460;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease S-5 precursor (EC 3.1.27.1) (S5-RNase).
OS Pyrus pyrifolia (Japanese pear) (Pyrus serotina).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside I; Rosales; Rosaceae; Maloideae; Pyrus.
OX NCBI_TaxID=3767;

RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=cv, Kousui; TISSUE=Style;
 RA Sasa H., Hirano H.;
 RT "Nucleotide sequence of a cDNA encoding 5S-Rnase from Japanese pear
 (Pyrus serotina).";
 RL (In) Plant Gene Register PGR97-007.
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=cv, Hosui; TISSUE=Style;
 RX MEDLINE=98363646; PubMed=9700066;
 RA Ishimizu T., Shinkawa T., Sakiyama F., Norioka S.;
 RT "Primary structural features of rosaceous S-Rnases associated with
 gametophytic self-incompatibility.";
 RL Plant Mol. Biol. 37:931-941(1998).
 RN [3]
 RP CARBOHYDRATE-LINKAGE SITES, AND STRUCTURE OF CARBOHYDRATES.
 RC STRAIN=cv, Hosui; TISSUE=Style;
 RX MEDLINE=99398327; PubMed=10469125;
 RA Ishimizu T., Mitsuoka Y., Shinkawa T., Natsuka S., Hase S.,
 RT "Presence of asparagine-linked N-acetylglucosamine and chitobiose in
 Pyrus pyrifolia S-Rnases associated with gametophytic self-
 incompatibility.";
 RL Eur. J. Biochem. 263:624-634(1999).
 CC -1- FUNCTION: SELF-INCOMPATIBILITY (SI) IS THE INHERITED ABILITY OF
 A FLOWERING PLANT TO PREVENT SELF-FERTILIZATION BY DISCRIMINATING
 BETWEEN SELF AND NON-SELF POLLEN DURING POLLINATION. IN MANY
 SPECIES, SELF-INCOMPATIBILITY IS CONTROLLED BY
 THE SINGLE, MULTIALLELIC LOCUS S.
 CC -1- CATALYTIC ACTIVITY: Two-stage endonucleolytic cleavage to
 nucleoside 3'-phosphates and 3'-phosphooligonucleotides with
 2',3'-cyclic phosphate intermediates.
 CC -1- PTM: N-Glycan AT ASN-45 CONSISTS OF DISACCHARIDE (GLCNAC-GLCNAC).
 CC -1- N-LINKED CORE STRUCTURE AT ASN-143 CONTAINS XYLOSE.
 CC -1- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, D88282; BAA13577.1;
 CC EMBL, AB002141; BAA32412.1;
 CC InterPro, IPR001568; RNase_T2.
 CC Pfam, PF00445; ribonuclease_T2.1.
 CC PROSITE, PS00530; RNase_T2_1; 1.
 DR KM Hydroxylase; Endonuclease; Glycoprotein; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 227
 FT ACT_SITE 60 60 RIBONUCLEASE S-5.
 FT ACT_SITE 111 111 BY SIMILARITY.
 FT ACT_SITE 115 115 BY SIMILARITY.
 FT ACT_SITE 115 115 BY SIMILARITY.
 FT DISULFID 75 118 BY SIMILARITY.
 FT DISULFID 182 208 BY SIMILARITY.
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 227 AA; 26059 MW; 44BC2A9E9E738FCA CRC64;
 SQ SEQUENCE
 Query Match 100.0%; Score 22; DB 1; Length 227;
 Best Local Similarity 100.0%; Pred. NO. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 19
 RNSI_PRRPY

ID RNSI_PRRPY STANDARD; PRT; 228 AA.
 AC 080322;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ribonuclease S-1 precursor (EC 3.1.27.1) (SI-RNase).
 OS Pyrus pyrifolia (Japanese pear) (Pyrus serotina).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids I; Rosales; Rosaceae; Maloideae; Pyrus.
 OC NCBI_taxid=3767;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=cv, Imamuraaki; TISSUE=Style;
 RX MEDLINE=98363646; PubMed=9700066;
 RA Ishimizu T., Shinkawa T., Sakiyama F., Norioka S.;
 RT "Primary structural features of rosaceous S-Rnases associated with
 gametophytic self-incompatibility.";
 RL Plant Mol. Biol. 37:931-941(1998).
 RN [2]
 RP CARBOHYDRATE-LINKAGE SITES, AND STRUCTURE OF CARBOHYDRATES.
 RC STRAIN=cv, Imamuraaki; TISSUE=Style;
 RX MEDLINE=99398327; PubMed=10469125;
 RA Ishimizu T., Mitsuoka Y., Shinkawa T., Natsuka S., Hase S.,
 RT "Presence of asparagine-linked N-acetylglucosamine and chitobiose in
 Pyrus pyrifolia S-Rnases associated with gametophytic self-
 incompatibility.";
 RL Eur. J. Biochem. 263:624-634(1999).
 CC -1- FUNCTION: SELF-INCOMPATIBILITY (SI) IS THE INHERITED ABILITY OF
 A FLOWERING PLANT TO PREVENT SELF-FERTILIZATION BY DISCRIMINATING
 BETWEEN SELF AND NON-SELF POLLEN DURING POLLINATION. IN MANY
 SPECIES, SELF-INCOMPATIBILITY IS CONTROLLED BY
 THE SINGLE, MULTIALLELIC LOCUS S (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Two-stage endonucleolytic cleavage to
 nucleoside 3'-phosphates and 3'-phosphooligonucleotides with
 2',3'-cyclic phosphate intermediates.
 CC -1- PTM: N-LINKED CORE STRUCTURE AT ASN-87 AND ASN-101 CONTAINS XYLOSE
 AND FUCOSE OR CONSISTS OF DISACCHARIDE (GLCNAC-GLCNAC). N-LINKED
 CORE STRUCTURE AT ASN-144 CONTAINS XYLOSE.
 CC -1- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, AB002139; BAA32412.1;
 CC InterPro, IPR001568; RNase_T2.
 CC Pfam, PF00445; ribonuclease_T2.1.
 CC PROSITE, PS00530; RNase_T2_1; 1.
 DR KM Hydroxylase; Nuclease; Endonuclease; Glycoprotein; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 228
 FT ACT_SITE 75 119 RIBONUCLEASE S-1.
 FT ACT_SITE 183 210 BY SIMILARITY.
 FT ACT_SITE 60 60 BY SIMILARITY.
 FT ACT_SITE 112 112 BY SIMILARITY.
 FT ACT_SITE 116 116 BY SIMILARITY.
 FT ACT_SITE 116 116 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 228 AA; 25791 MW; 95AF4EB28B6D5718 CRC64;
 SQ SEQUENCE
 Query Match 100.0%; Score 22; DB 1; Length 228;
 Best Local Similarity 100.0%; Pred. NO. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 DB 35 QOYO 38

RESULT 20
 RNS4_PYPY STANDARD; PRT; 228 AA.

AC 040966;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ribonuclease S-4 precursor (EC 3.1.27.1) (S4-RNase).
 OS *Pyrus pyrifolia* (Japanese pear) (*Pyrus serotina*).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosidia I; Rosales; Rosaceae; Maloideae; Pyrus.
 NCBI_TaxID=3767;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nijisseiki;
 RA MEDLINE=98241513; PubMed=9573352;
 RT Ushijima K., Sase H., Hirano H.;
 RT "Characterization of the flanking regions of the S-RNase genes of
 RT Japanese pear (*Pyrus serotina*) and apple (*Malus domestica*).";
 RL Gene 211:159-167(1998).
 [2]
 RP SEQUENCE OF 6-228 FROM N.A.
 RC STRAIN=cv. Nijisseiki; TISSUE=Style;
 RA Noriooka N., Ohnishi Y., Noriooka S., Ishimizu T., Nakanishi T.,
 RA Sakiyama F.;
 RT "Nucleotide sequences of cDNAs encoding S2- and S4-RNases from
 RT Japanese pear (*Pyrus pyrifolia* Nakai 1).";
 RL (in) Plant Gene Register PGR95-020.
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nijisseiki;
 RA Noriooka N., Noriooka S.;
 RT "The genomic structures of S-RNases in Japanese pear.";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=cv. Nijisseiki;
 RA MEDLINE=97044748; PubMed=8889819;
 RA Noriooka N., Noriooka S., Ohnishi Y., Ishimizu T., Oneyama C.,
 RA Nakanishi T., Sakiyama F.;
 RT "Molecular cloning and nucleotide sequences of cDNAs encoding S-allele
 RT specific styelar RNases in a self-incompatible cultivar and its self-
 RT compatible mutant of Japanese pear, *Pyrus pyrifolia* Nakai.";
 RL J. Biochem. 120:335-345(1996).
 [5]
 RP CARBOHYDRATE-LINKAGE SITES, AND STRUCTURE OF CARBOHYDRATES.
 RC STRAIN=cv. Nijisseiki; TISSUE=Style;
 RA MEDLINE=99398327; PubMed=10469125;
 RA Ishimizu T., Mitsuoka Y., Shinkawa T., Natsuka S., Hase S.,
 RA Miyagi M., Sakiyama F., Noriooka S.;
 RT "Presence of asparagine-linked N-acetylglucosamine and chitobiose in
 RT *Pyrus pyrifolia* S-RNases associated with gametophytic self-
 RT incompatibility.";
 RL Eur. J. Biochem. 263:624-634(1999).
 [1]
 RP FUNCTION: SELF-INCOMPATIBILITY (SI) IS THE INHERITED ABILITY OF
 RP A FLOWERING PLANT TO PREVENT SELF-FERTILIZATION BY DISCRIMINATING
 RP BETWEEN SELF AND NON-SELF POLLEN DURING POLLINATION. IN MANY
 RP SPECIES, SELF-INCOMPATIBILITY IS CONTROLLED BY
 RP THE SINGLE MULTIALLELIC LOCUS S.
 CC -1- CATALYTIC ACTIVITY: Two-stage endonucleolytic cleavage to
 CC nucleoside 3'-phosphates and 3'-phosphonucleosides with
 CC 2',3'-cyclic phosphate intermediates.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- PTM: N-GLYCAN AT ASN-101, ASN-160, AND ASN-175 CONSISTS OF
 CC DISACCHARIDE (GLCNAc-GLCNAc). N-LINKED CORE STRUCTURE AT ASN-144
 CC CONTAINS XYLOSE.
 CC -1- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; AB009385; BAA28354.1; -
 DR EMBL; D49528; BAA08474.1; -
 DR EMBL; AB014072; BAA77692.1; -
 DR PIR; JC4869; JC4869.
 DR InterPro; IPR001568; RNase T2.
 DR Pfam; PF00445; ribonuclease_T2; 1.
 DR PROSITE; PS00530; RNASE_T2_1; 1.
 KM Hydroxylase; Nuclease; Endonuclease; Glycoprotein; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 228 RIBONUCLEASE S-4.
 FT ACT SITE 60 60 BY SIMILARITY.
 FT ACT_SITE 112 112 BY SIMILARITY.
 FT ACT_SITE 116 116 BY SIMILARITY.
 FT DISULFID 75 119 BY SIMILARITY.
 FT DISULFID 183 210 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAc).
 FT CARBOHYD 101 101 N-LINKED (GLCNAc).
 FT CARBOHYD 144 144 N-LINKED (GLCNAc).
 FT CARBOHYD 160 160 N-LINKED (GLCNAc).
 FT CARBOHYD 175 175 N-LINKED (GLCNAc).
 SQ SEQUENCE 228 AA; 25858 MW; C8A95158FF1A/CDD CRC64;

Query Match 100.0%; Score 22; DB 1; Length 228;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 DB 35 QOYO 38

RESULT 21
 RNS6_PYPY STANDARD; PRT; 229 AA.

AC 080324;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Ribonuclease S-6 precursor (EC 3.1.27.1) (S6-RNase).
 OS *Pyrus pyrifolia* (Japanese pear) (*Pyrus serotina*).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosidia I; Rosales; Rosaceae; Maloideae; Pyrus.
 NCBI_TaxID=3767;
 RP [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=cv. Imamuraaki; TISSUE=Style;
 RA MEDLINE=98363646; PubMed=9700066;
 RA Ishimizu T., Shinkawa T., Sakiyama F., Noriooka S.;
 RT "Primary structural features of roseaceous S-RNases associated with
 RT gametophytic self-incompatibility.";
 RL Plant Mol. Biol. 37:931-941(1998).
 [2]
 RP CARBOHYDRATE-LINKAGE SITES, AND STRUCTURE OF CARBOHYDRATES.
 RC STRAIN=cv. Imamuraaki; TISSUE=Style;
 RA MEDLINE=99398327; PubMed=10469125;
 RA Ishimizu T., Mitsuoka Y., Shinkawa T., Natsuka S., Hase S.,
 RA Miyagi M., Sakiyama F., Noriooka S.;
 RT "Presence of asparagine-linked N-acetylglucosamine and chitobiose in
 RT *Pyrus pyrifolia* S-RNases associated with gametophytic self-
 RT incompatibility.";
 RL Eur. J. Biochem. 263:624-634(1999).
 [1]
 RP FUNCTION: SELF-INCOMPATIBILITY (SI) IS THE INHERITED ABILITY OF
 RP A FLOWERING PLANT TO PREVENT SELF-FERTILIZATION BY DISCRIMINATING

BETWEEN SELF AND NON-SELF POLLEN DURING POLLINATION. IN MANY SPECIES, SELF-INCOMPATIBILITY IS CONTROLLED BY THE SINGLE, MULTIALLELIC LOCUS S.

-1- CATALYTIC ACTIVITY: Two-stage endonucleolytic cleavage to nucleoside 3'-phosphates and 3'-phosphooligonucleotides with 2',3'-cyclic phosphate intermediates.

-1- PFM: N-GLYCINYL AT ASN-188 AND ASN-203 CONSISTS OF MONOSACCHARIDE (GLCNAC) OR DISACCHARIDE (GLCNAC-GLCNAC).

-1- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB002142; BAA32415.1; -
InterPro; IPR001568; RNase T2.
Pfam; PF00445; ribonuclease_T2; 1.
DR PROSITE; PS00530; RNASE T2_1; 1.
KM Hydrolase; Nuclease; Glycoprotein; Signal.
FT SIGNAL 1 27
FT CHAIN 28 229
FT ACT_SITE 60 60 RIBONUCLEASE S-6.
FT ACT_SITE 112 112 BY SIMILARITY.
FT ACT_SITE 116 116 BY SIMILARITY.
FT DISULFID 75 119 BY SIMILARITY.
FT DISULFID 184 210 BY SIMILARITY.
FT CARBOHYD 77 77 N-LINKED (GLCNAC).
FT CARBOHYD 87 87 N-LINKED (GLCNAC).
FT CARBOHYD 145 145 N-LINKED (GLCNAC OR GLCNAC). (HIGH MANNOSE).
FT CARBOHYD 188 188 N-LINKED (GLCNAC OR GLCNAC).
FT CARBOHYD 203 203 N-LINKED (GLCNAC OR GLCNAC).
SQ SEQUENCE 229 AA; 26278 MW; 907C5231E0B612A3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 22; DB 1; Length 229;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
Db 35 QOYQ 38

UNT 22
7_BPP22 STANDARD; PRT; 229 AA.
Q01074;
01-OCT-1993 (Rel. 27, Created)
01-OCT-1993 (Rel. 27, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA transfer protein gp7 precursor.
GN
OS Bacteriophage P22.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC P22-like viruses.
OX NCBI_Taxid=10754;
RN
RN SEQUENCE FROM N.A.
RX MEDLINE=92394890; PubMed=1522065;
RA Conlin C.A., Vism B.R., Miller C.G.;
RT "Oligopeptidase A is required for normal phage P22 development.";
RN J. Bacteriol. 174:5869-5880(1992).
RN
RN SEQUENCE FROM N.A.
RA Kropinski A.M.B., Vanderbyl C.S.;
RT "The completed sequence of genome of Salmonella phage P22.";
RN Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
RN
RN SEQUENCE OF 225-229 FROM N.A.
RX MEDLINE=93219140; PubMed=8464750;

Adhikari P., Berger P.B.;
RT "Sequence of a DNA injection gene from Salmonella typhimurium phage P22.";
RN Nucleic Acids Res. 21:1499-1499(1993).
RN (4)
RP SEQUENCE OF 21-31.
RX MEDLINE=91306435; PubMed=1853558;
RA Eppler K., Wyckoff E., Goates J., Parr R., Casjens S.;
RT "Nucleotide sequence of the bacteriophage P22 genes required for DNA packaging";
RN Virology 183:519-538(1991).
RL
CC -1- FUNCTION: REQUIRED FOR EJECTION OF THE PHAGE DNA FROM THE PHAGE PARTICLE AND INJECTION OF THAT DNA INTO THE HOST.
CC -1- SIMILARITY: STRONG, TO PHAGE APSE-1 P32.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M93985; AAA72115.1; -
EMBL; AF217253; AAF75053.1; -
DR EMBL; L07556; -; NOT_ANNOTATED_CDS.
DR PIR; C43330; C43330.
KM Late protein.
FT PROPEP 1 20 REMOVED IN MATURE FORM.
FT CHAIN 21 229 DNA TRANSFER PROTEIN GP7.
SQ SEQUENCE 229 AA; 23407 MW; 051A26642P185982 CRC64;

Query Match
Best Local Similarity 100.0%; Score 22; DB 1; Length 229;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
Db 90 QOYQ 93

RESULT 23
TBB3 ANEPH STANDARD; PRT; 239 AA.
AC P33632;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tubulin beta-3 chain (Fragment).
GN TUBB3.
OS Anemita phyllitidis (Fern).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Filicophyta; Filicopsida; Filicales; Schizaeaceae; Anemita.
OX NCBI_Taxid=12940;
RN
RN SEQUENCE FROM N.A.
RA Moeppe B., Maucher H.P., Bogenberger J.M., Schrandolf H.;
RT Submitted (APR-1993) to the EMBL/Genbank/DBJ databases.
RL
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: Dimer of alpha and beta chains.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X69187; CAA48931.1; ALT_INIT.

DR PIR, S32670; S32670.
 DR InterPro: IPR002453; Beta tubulin.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin_1.
 DR Pfam: PF03953; tubulin_C_1.
 DR PROSITE: PS00227; TUBULIN, PARTIAL.
 DR PROSITE: PS00228; TUBULIN, B. AUTOREG; PARTIAL.
 DR Microtubules; GTP-binding; Multigene family.
 KM NON TER 1 1
 FT SEQUENCE 239 AA; 27436 MW; EF4E367CF11EB16C CRC64;
 SQ
 Query Match 100.0%; Score 22; DB 1; Length 239;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYQ 4
 DB 219 QOYQ 222
 RESULT 24
 DABP_STAM STANDARD; PRT; 240 AA.
 ID DABP_STAM STANDARD; PRT; 240 AA.
 AC Q99088;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Dihydrodipicolinate reductase (EC 1.3.1.26) (DHPR).
 GN DABP OR SAV1396 OR SA1228.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 RX NCBI_TaxID=158878, 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mu50 / ATCC 700699, and N315;
 RX MEDLINE=2111952; PubMed=1181816;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iino J.-O., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
 RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshida K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 aureus.";
 RL Lancet 357:1225-1240(2001).
 CC -1- CATALYTIC ACTIVITY: 2,3,4,5-tetrahydrodipicolinate + NAD(P)(+) =
 2,3-dihydrodipicolinate + NAD(P)H.
 CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
 semialdehyde; second step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the dihydrodipicolinate reductase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL, AP003362; BAB57558.1; -
 DR EMBL, AP003133; BAB42488.1; -
 DR PIR, D89916; D89916.
 DR HAMAP, MF 00102; -; 1.
 DR HAMAP, PF04036; 1DRM.
 DR InterPro: IPR000846; DABP.
 DR InterPro: IPR000846; DABP.
 DR Pfam: PF01113; DABP_N; 1.
 DR Pfam: PF01113; DABP_N; 1.
 DR ProDom: PD004105; DABP; 1.
 DR ProDom: PD004105; DABP; 1.
 DR TIGRFAMs, TIGR00036; dapB; 1.
 DR TIGRFAMs, TIGR00036; dapB; 1.

DR PROSITE, PS01298; DABP; FALSE NEG.
 KM Diaminopimelate biosynthesis; Lysine biosynthesis; Oxidoreductase;
 KM NADP. Complete proteome.
 KM NADP. Complete proteome.
 SQ SEQUENCE 240 AA; 26726 MW; F910FBD3DBFB8D CRC64;
 Query Match 100.0%; Score 22; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYQ 4
 DB 41 QOYQ 44
 RESULT 25
 DABP_STAM STANDARD; PRT; 240 AA.
 ID DABP_STAM STANDARD; PRT; 240 AA.
 AC Q9E211;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Dihydrodipicolinate reductase (EC 1.3.1.26) (DHPR).
 GN DABP.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 RX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8325-4;
 RA Wilshire M.D., Burnham M.K., Foster S.J.;
 RT "Identification and analysis of Staphylococcus aureus components
 RT involved during growth in sera."
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: 2,3,4,5-tetrahydrodipicolinate + NAD(P)(+) =
 2,3-dihydrodipicolinate + NAD(P)H.
 CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
 semialdehyde; second step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the dihydrodipicolinate reductase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL, AF306669; AAG42247.1; ALT_INIT.
 DR HAMAP, MF 00102; -; 1.
 DR HAMAP, PF04036; 1DRM.
 DR InterPro: IPR000846; DABP.
 DR Pfam: PF05173; DABP_C; 1.
 DR Pfam: PF01113; DABP_N; 1.
 DR ProDom: PD004105; DABP; 1.
 DR TIGRFAMs, TIGR00036; dapB; 1.
 DR PROSITE, PS01298; DABP; FALSE NEG.
 KM Diaminopimelate biosynthesis; Lysine biosynthesis; Oxidoreductase;
 KM NADP.
 SQ SEQUENCE 240 AA; 26583 MW; 230512D0F7588D8D CRC64;
 Query Match 100.0%; Score 22; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYQ 4
 DB 41 QOYQ 44
 RESULT 26
 DABP_STAM STANDARD; PRT; 240 AA.
 ID DABP_STAM STANDARD; PRT; 240 AA.

AC Q8NWS4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Dihydrodipicolinate reductase (EC 1.3.1.26) (DHRP)
 GN DABP OR MW1284.
 OS Staphylococcus aureus (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwana N., Asano K., Naimi T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramoto K.;
 RT "Genome and virulence determinants of high virulence community-
 RT acquired MRSA.";
 RT Lancet 359:1819-1827(2002).
 CC -1- CATALYTIC ACTIVITY: 2,3,4,5-tetrahydrodipicolinate + NAD(P)(+) =
 CC 2,3-dihydrodipicolinate + NAD(P)H.
 CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
 CC semialdehyde; second step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the dihydrodipicolinate reductase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP004826; BAB95149.1; -
 DR HAMAP, MF_00102; 1.
 DR InterPro: IPR000846; DABP.
 DR Pfam: PF05173; DABP_C; 1.
 DR Pfam: PF01113; DABP_N; 1.
 DR ProDom: PD004105; DABP_N; 1.
 DR TIGRFAMs: TIGR00036; dabp; 1.
 DR PROSITE: PS01298; DABP; FAUSE_NEG.
 DR Diaminopimelate biosynthesis; Lysine biosynthesis; Oxidoreductase;
 KM NADP; Complete proteome.
 SQ SEQUENCE 240 AA; 26722 MW; B87D3B14C7645299 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 22; DB 1; Length 240;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 DB 41 QOYQ 44

RESULT 27
 YB46_PASMU STANDARD; PRT; 241 AA.
 AC P57911;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein PM1146.
 GN PM1146.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.U., Zhang Q., Li L.U., Paustian M.L., Whitam T.S., Kapur V.,
 RA "Complete genomic sequence of Pasteurella multocida PM70.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 CC -1- SIMILARITY: BELONGS TO THE UPF0100 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB006155; AAK03230.1; -
 DR HAMAP, MF_00035; atypical; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 241 AA; 27100 MW; 466918D8A94B10B CRC64;

Query Match
 Best Local Similarity 100.0%; Score 22; DB 1; Length 241;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 DB 22 QOYQ 25

RESULT 28
 P64L_DROME
 ID P64L_DROME STANDARD; PRT; 251 AA.
 AC Q9VA12; Q81S89;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Proteasome subunit alpha type 4-like (EC 3.4.25.1).
 GN PROS-ALPHA-3T OR CG1376.
 OS Drosophila melanogaster (Fruit Fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
 RC TISSUE=Testis;
 RX PubMed=12421421;
 RX Ma J., Katz E., Belote J.M.;
 RT "Expression of proteasome subunit isoforms during spermatogenesis in
 RT Drosophila melanogaster.";
 RL Insect Mol. Biol. 11:627-639(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoeklin R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport I.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.R., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Morkulov G., Mishina N.V., Mobarry C., Morris J., Moshneff A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
 RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Ghibb R.A., Myers B.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: The proteasome is a multicatalytic proteinase complex
 CC which is characterized by its ability to cleave peptides with Arg,
 CC Phe, Tyr, Leu, and Glu adjacent to the leaving group at neutral or
 CC slightly basic pH. The proteasome has an ATP-dependent proteolytic
 CC activity.
 CC -1- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
 CC specificity.
 CC -1- PATHWAY: Involved in an ATP/ubiquitin-dependent non-lysosomal
 CC proteolytic pathway.
 CC -1- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
 CC SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Testis, prominent after meiosis II. After
 CC meiosis, predominantly localized to the haploid spermatid nuclei
 CC of the 64-cell cysts, remaining during the elongation and
 CC condensation of the spermatid nuclei. In mature, motile sperm,
 CC expression is seen exclusively in the sperm head.
 CC -1- SIMILARITY: Belongs to peptidase family 11A.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: AE003776; AAF57116.1; -;
 CC DR EMBL: AY147240; AAN63094.1; -;
 CC DR HSSP: P23638; IIRP.
 CC DR MEROPS: T01.973; -;
 CC DR FLYBase: FBgn0039819; Pros-alpha-3T.
 CC DR InterPro: IPR000426; Proteasome A.
 CC DR InterPro: IPR001353; Proteasome.
 CC DR Pfam: PF00227; proteasome; 1.
 CC DR PROSITE: PS00388; PROTEASOME_A; 1.
 CC DR KMW: Hydrolyase; Protease; Nuclear protein; Proteasome.
 CC FT CONFLICT 74 CA -> WP (IN REF. 1).
 CC FT CONFLICT 133 F -> L (IN REF. 1).
 CC FT CONFLICT 141 RF -> AS (IN REF. 1).
 CC SO SEQUENCE 251 AA; 28169 MW; BIAFP83978BD1d3 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 251;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 DB 94 QOYO 97

RESULT 29
 NEP_HV2B STANDARD; PRT; 257 AA.
 AC P18092;
 DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Negative factor (P-protein) (27 kDa protein) (3'ORF).
 GN NEF.
 OS Human immunodeficiency virus type 2 (isolate BEN) (HIV-2).
 OC Viruses; Retroid viruses; Retroviridae; Lentiviruses.
 ON NCBI_Taxid=11714;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90281594; PubMed=2353457;
 RA Kirchhoff F., Jentsch K., Bachmann B., Stuke A., Laloux C.,
 RA Lueke W., Stahl-Hennig C., Schneider J., Nieselt K., Eigen M.,
 RA Humann G.;
 RT "A novel proviral clone of HIV-2: biological and phylogenetic
 RT relationship to other primate immunodeficiency viruses.";
 RL Virology 177:305-311(1990).
 CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
 CC activities. It seems to down-regulate the CD4(74) antigen.
 CC -1- MISCELLANEOUS: THIS ISOLATE IS FROM A GERMAN AIDS PATIENT (WITH
 CC PREDOMINANTLY NEUROLOGICAL COMPLICATIONS) WHO WAS PROBABLY
 CC INFECTED IN MALI.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: M30502; AAB00744.1; -;
 CC DR HSSP: P03406; IERN.
 CC DR HIV: M30502; NEF52BEN.
 CC DR InterPro: IPR001558; HIV_Nef.
 CC DR Pfam: PF00469; F-protein; 1.
 CC DR ProDom: PD000031; HIV_Nef; 1.
 CC DR AIDS; Myristate; GTP-binding.
 CC FT LIPID 2
 CC FT LIPID 2
 CC SO SEQUENCE 257 AA; 29911 MW; B8F8B5B21A813A1 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 257;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 DB 58 QOYO 61

RESULT 30
 NEP_HV2D1 STANDARD; PRT; 257 AA.
 AC P17753;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Negative factor (P-protein) (27 kDa protein) (3'ORF).
 GN NEF.
 OS Human immunodeficiency virus type 2 (isolate D194) (HIV-2).
 OC Viruses; Retroid viruses; Retroviridae; Lentiviruses.
 ON NCBI_Taxid=11713;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89184631; PubMed=2467304;
 RA Kuenzel H., von Briesen H., Dietrich U., Adamski M., Mix D.,
 RA Biesecl L., Kreutz R., Immlmann A., Henco K., Weichner C.,
 RA Andreesen R., Gelderblom H., Ruebsamen-Waigmann H.;
 RT "Molecular cloning of two west African human immunodeficiency virus
 RT type 2 isolates that replicate well in macrophages: a Gambian
 RT isolate, from a patient with neurologic acquired immunodeficiency
 RT syndrome, and a highly divergent Ghanaian isolate.";
 RT Proc. Natl. Acad. Sci. U.S.A. 86:2383-2387(1989).
 RL [2]


```

RP SEQUENCE FROM N.A.
RA MEDLINE=91045094; PubMed=2235509;
RX Kuehnelt H., Kreutz R., Ruebsamen-Waigmann H.;
RT "Nucleotide sequence of HIV-2D194, an isolate from a Gambian case of
RL 'neuro-AIDS', which showed excellent growth in macrophages.";
CC Nucleic Acids Res. 18:6142-6142(1990).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(74) antigen.
CC -1- MISCELLANEOUS: THIS ISOLATE IS FROM A GAMBIAN CASE OF
CC 'NEURO-AIDS'.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; J04542; AAA76848.1; -
EMBL; X52223; CAA36472.1; -
DR PIR; S12160; S12160.
DR HIV; J04542; NEF2D194.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
DR KW; Myristate; GTP-binding.
DR LIPID 2
DR FT SEQUENCE 257 AA; 29449 MW; 2C5B1B5C8ABCF3 CRC64;
SQ
Query Match 100.0%; Score 22; DB 1; Length 257;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYQ 4
DB 58 QOYQ 61
RESULT 31
TR33_YEAST STANDARD; PRT; 268 AA.
ID TR33_YEAST
AC Q99394;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
TR33 OR YOR115C OR O3251 OR YOR3251C.
Transport protein particle 33 kDa subunit (TRAPP 33 kDa subunit).
Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxId=4932;
[1]
SEQUENCE FROM N.A.
RA STRAIN=S288C / FY1679;
RX MEDLINE=97060020; PubMed=8904341;
RA Wiemann S., Rechmann S., Benes V., Voes H., Schwager C., Vleck C.,
RA Stegemann U., Zimmermann U., Erfle H., Pece V., Ansoerge W.;
RT "Sequencing and analysis of 51 kb on the right arm of chromosome XV
RL from Saccharomyces cerevisiae reveals 30 open reading frames.";
RA Yeast 12:281-288(1996).
[2]
SEQUENCE FROM N.A.
RA MEDLINE=97344368; PubMed=9200815;
RA Voes H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoru C.,
RA Schwager C., Paces V., Sander C., Ansoerge W.;
RT "DNA sequencing and analysis of 130 kb from yeast chromosome XV,
RL Yeast 13:655-672(1997).
-1- FUNCTION: TRAPP plays a key role in the late stages of endoplasmic
reticulum to Golgi traffic.
-1- SUBUNIT: Part of multisubunit complex composed of BET3, BET5,
TR320, TR331, TR333, TR365, TR385, TR3120 and TR3130.
-1- SUBCELLULAR LOCATION: Cis-Golgi complex.

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X90518; CAA62124.1; -
DR EMBL; X94335; CAA64035.1; -
DR EMBL; Z75023; CAA9313.1; -
DR PIR; S61003; S61003.
DR SGD; S0005641; TR333.
DR GO; GO:0030008; C-TRAPP; IDA.
DR GO; GO:0006888; P:ER to Golgi transport; IMP.
DR KW; Transport; Endoplasmic reticulum; Golgi stack.
DR SQ SEQUENCE 268 AA; 30749 MW; 2480F69D7A48D363 CRC64;
SQ
Query Match 100.0%; Score 22; DB 1; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYQ 4
DB 28 QOYQ 31
RESULT 32
ABCD_HUMAN STANDARD; PRT; 274 AA.
ID ABCD_HUMAN
AC Q9NSE7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Putative ATP-binding cassette transporter C13 (Protein PREB6).
GN ABCC13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
[1]
SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Liver, and Placenta;
RA Yabuchi H., Takayanagi S.-I., Ishikawa T.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Matanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishi K., Totoki Y., Choi D.-K., Groner Y.,
RA Soeda E., Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordliek G., Hornischer K., Brandt P.,
RA Scharfe W., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramer J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Weinmeyer S., Bozzym K., Gardner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RA Nature 405:311-319(2000).
[3]
TISSUE SPECIFICITY.
RA MEDLINE=22032985; PubMed=12036298;
RA Gardiner K., Slavov D., Bechtel L., Davison M.;
RT "Annotation of human chromosome 21 for relevance to down syndrome:
RL gene structure and expression analysis.";
CC Genomics 79:833-843(2002).
-1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named Isoforms=3;
CC Comment=Experimental confirmation may be lacking for some

```

```

CC isoforms;
CC Name1; Synonym=A;
CC IsoId=Q9NSE7-1; Sequence=Displayed;
CC Name2;
CC IsoId=Q9NSE7-2; Sequence=VSP_000066;
CC Name3; Synonym=B;
CC IsoId=Q9NSE7-3; Sequence=VSP_000067.
CC -1- TISSUE SPECIFICITY: Prefrontal cortex and fetal tissues.
CC -1- SIMILARITY: TO THE N-TERMINAL PART OF ABC TRANSPORTERS FROM THE
CC ABC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY063514; AAL47201.1; -.
CC EMBL; AF418600; AAL99903.1; -.
CC EMBL; AY063515; AAL47202.1; -.
CC EMBL; AP001660; BAA95546.1; -.
CC Gene; HGNC:16022; ABCC13.
CC InterPro; IPR001140; ABC_TM_transp.
CC Pfam; PF00664; ABC_membrane_1.
CC Transmembrane; Alternative splicing.
CC TRANSMEM 118 138
CC TRANSMEM 154 174
CC VARSPLEM 200 274
CC -----
CC ALLSVNRQKSTGTEIINLSATGIDSKSPQSLVCPSPN
CC PNCRIPLARAGSSVSRRGSPCVCTNKCPCSN -> VSL
CC ATLCVFLDEGNILATKVTSMLEILRILPLFLPYI
CC SAVVOTKISLGRLEDPLNTELLPOSILENTYGTDAIGFD
CC ASFSMKTGMPULKEALMLMFLSRPRLAFCCKTSLAPS
CC (in isoform 2).
CC /FTId=VSP_000066.
CC /FTId=VSP_000067.
CC /FTId=VSP_000067.
CC Missing (in isoform 3).
CC -----
CC VARSPLEM 141 274
CC SEQUENCE 274 AA; 30831 MW; 02BDFCFE0FED623F CRC64;
CC -----
CC Query Match 100.0%; Score 22; DB 1; Length 274;
CC Best Local Similarity 100.0%; Pred. No. 1.3e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 1 QOYQ 4
CC ||||
CC 175 QOYQ 178
CC -----
CC ULT 33
CC EMBL
CC PYRF EMBL STANDARD; PRT; 274 AA.
CC -----
CC 01-JUL-1989 (Rel. 11, Created)
CC 01-JUL-1989 (Rel. 11, Last sequence update)
CC 28-FEB-2003 (Rel. 41, Last annotation update)
CC Oxidative 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
CC (OMPCase) (OMPDecase) (Uridine 5'-monophosphate synthase) (UMP
CC synthase).
CC DE
CC GN
CC OS Emericella nidulans (Aspergillus nidulans).
CC OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
CC Eucotiiales; Trichocomaceae; Emericella.
CC NCBI_TaxID=162425;
CC -----
CC SEQUENCE FROM N.A.
CC MEDLINE=88185841; PubMed=3328733;
CC Oakley B.R., Rinehart J.B., Mitchell B.L., Oakley C.E., Carmona C.,
CC Gray G.L., May G.S.;
CC "Cloning, mapping and molecular analysis of the pyrG (orotidine-5'-
CC phosphate decarboxylase) gene of Aspergillus nidulans.";
CC Gene 61:385-399(1987).
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate = UMP + CO(2).

```

```

CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19132; AAB6359.1; -.
CC DR PIR; A29630; DCA50E.
CC DR HSSP; P03962; IDOW.
CC DR InterPro; IPR001754; OMPDecase.
CC Pfam; PF00215; OMPDecase; 1.
CC PROSITE; PS00156; OMPDECASE; 1.
CC DR Pyrimidine biosynthesis; Lyase; Decarboxylase.
CC ACT SITE 95 95
CC BY SIMILARITY.
CC SEQUENCE 274 AA; 30053 MW; 485C671BFE47D6B CRC64;
CC -----
CC Query Match 100.0%; Score 22; DB 1; Length 274;
CC Best Local Similarity 100.0%; Pred. No. 1.3e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 1 QOYQ 4
CC ||||
CC 225 QOYQ 228
CC -----
CC RESULT 34
CC PYRF ASPNG STANDARD; PRT; 277 AA.
CC ID P07817; Q9HGS5;
CC 01-AUG-1988 (Rel. 08, Created)
CC 01-AUG-1988 (Rel. 08, Last sequence update)
CC 28-FEB-2003 (Rel. 41, Last annotation update)
CC Oxidative 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
CC (OMPDecase) (OMPDecase) (Uridine 5'-monophosphate synthase) (UMP
CC synthase).
CC DE
CC GN
CC OS Aspergillus niger.
CC OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
CC Eucotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
CC NCBI_TaxID=5061;
CC -----
CC SEQUENCE FROM N.A.
CC MEDLINE=88185829; PubMed=3357784;
CC Wilson L.J., Ward M., Carmona C.;
CC "Sequence of the Aspergillus niger pyrG gene.";
CC Nucleic Acids Res. 16:2339-2339(1988).
CC -----
CC SEQUENCE FROM N.A.
CC STRAIN=U112;
CC MEDLINE=88185829; PubMed=3357784;
CC van den Hombergh J.P.T.W., de Vries R.P., de Zwart P.J.I.,
CC "The pyrG gene from Aspergillus niger: characterization of gene and
CC mutants; targeting to homologous locus and use of disruption strains
CC in locus-specific integration.";
CC Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate = UMP + CO(2).
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; X06626; CAA29838.1; -
DR EMBL; X96734; CAA65508.2; -
DR PIR; S03652; DCA50N.
DR HSSP; P03962; IDOW.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase.1.
DR PROSITE; PS00156; OMPDECASE; 1.
KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
FT ACT SITE 95
FT VARIANT 73 D -> N (IN STRAIN N400).
FT VARIANT 109 R -> G (IN STRAIN N400).
FT VARIANT 145 S -> A (IN STRAIN N400).
FT VARIANT 192 S -> A (IN STRAIN N400).
FT VARIANT 233 A -> G (IN STRAIN N400).
SQ SEQUENCE 277 AA; 30196 MW; 02F30DAF5D9DA808 CRC64;

Query Match
Best Local Similarity 100.0%; Score 22; DB 1; Length 277;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYQ 4
227 QOYQ 230

RESULT 35
PYRF_ASPOR
ID PYRF_ASPOR STANDARD; PRT; 277 AA.
AC 013416;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
DE (OMPDCase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP
DE synthase).
GN PYRG.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocommataceae; mitosporic Trichocommataceae; Aspergillus.
RN NCBI_TaxID=5062;
RX [1]
RP SEQUENCE FROM N.A.
RA STRAIN=TK3;
RC Donnas A.; Vandenbroek P.J.M.; Affolter M.; Monod M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBA databases.
KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
FT ACT SITE 95
FT VARIANT 95 BY SIMILARITY.
SQ SEQUENCE 277 AA; 30262 MW; C6F2EB037B112977 CRC64;

Query Match
Best Local Similarity 100.0%; Score 22; DB 1; Length 277;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYQ 4
227 QOYQ 230

```

```

RESULT 36
PYRF_ASPFU
ID PYRF_ASPFU STANDARD; PRT; 278 AA.
AC 013410;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
DE (OMPDCase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP
DE synthase).
GN PYRG.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocommataceae; mitosporic Trichocommataceae; Aspergillus.
RN NCBI_TaxID=5085;
RX [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NIH 5233;
RC MEDLINE=98283859; PubMed=9618589;
RA Weidner G.; D'Enfert C.; Koch A.; Mol P.C.; Brakhage A.A.;
RT "Development of a homologous transformation system for the human
RT pathogenic fungus Aspergillus fumigatus based on the pyrg gene."
RT Curr. Genet. 33:378-385(1998).
RL Curr. Genet. 33:378-385(1998).
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate = UMP + CO(2).
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
DR EMBL; Y11303; CAA72161.1; -
DR HSSP; P03962; IDOW.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase.1.
DR PROSITE; PS00156; OMPDECASE; 1.
KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
FT ACT SITE 95
FT VARIANT 95 BY SIMILARITY.
SQ SEQUENCE 278 AA; 30267 MW; F4CF6703CFD0803 CRC64;

Query Match
Best Local Similarity 100.0%; Score 22; DB 1; Length 278;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QOYQ 4
227 QOYQ 230

RESULT 37
PYRF_CLAFU
ID PYRF_CLAFU STANDARD; PRT; 278 AA.
AC Q9HFV8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
DE (OMPDCase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP
DE synthase).
GN PYR1.
OS Cladosporium fulvum (Fulvia fulva).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.
RN NCBI_TaxID=5499;
RX [1]

```

RP SEQUENCE FROM N.A.
RA Sneljers S.S.;
RL "The pyr1 gene of the tomato pathogen Cladosporium fulvum";
Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Oxidoline 5'-phosphate = UMP + CO(2).
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF288696; MAG10516.1; -.
DR HSSP; P03962; IDOW.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase; 1.
DR PROSITE; PS00156; OMPDECASE; 1.
DR Pyrimidine biosynthesis; Lyase; Decarboxylase.
DR ACT_SITE 95
DR ACT_SITE 95 BY SIMILARITY.
SQ SEQUENCE 278 AA; 29956 MW; 6DDDB6B102127C63 CRC64;
Query Match 100.0%; Score 22; DB 1; Length 278;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
DB 226 QOYO 229
RESULT 39
ID PYR_FENCH STANDARD; PRT; 278 AA.
AC P09463;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oxidoline 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
DE (OMPdecase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP
DE synthase).
GN PYRG.
OS Penicillium chrysogenum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OC NCBI_TaxID=5076;
OC [1]
OC SEQUENCE FROM N.A.
OC STRAIN=AS-P-78;
RX MEDLINE=8835559; PubMed=3138658;
RA Cantoral J.M., Barredo J.L., Alvarez E., Diez B., Martin J.F.;
RT "Nucleotide sequence of the penicillium chrysogenum pyrG
RT (oxidoline-5'-phosphate decarboxylase) gene";
RL Nucleic Acids Res. 16:8177-8177(1988).
CC -1- CATALYTIC ACTIVITY: Oxidoline 5'-phosphate = UMP + CO(2).
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X08037; CA30835.1; -.
DR PIR; S01287; DCPLOC.
DR HSSP; P03962; IDOW.
DR InterPro; IPR001754; OMPdecase.

DR Pfam; PF00215; OMPdecase; 1.
DR PROSITE; PS00156; OMPDECASE; 1.
KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
FT ACT_SITE 98
FT ACT_SITE 98 BY SIMILARITY.
SQ SEQUENCE 278 AA; 29999 MW; 3BE30ED80DD209F8 CRC64;
Query Match 100.0%; Score 22; DB 1; Length 278;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
DB 228 QOYO 231
RESULT 39
ID PYR_SOLCC STANDARD; PRT; 280 AA.
AC Q9HF68;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oxidoline 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
DE (OMPdecase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP
DE synthase).
GN PYRG.
OS Solorina crocea.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Lecanoromycetes;
OC Lecanorales; Peltigerineae; Peltigeraceae; Solorina.
OC NCBI_TaxID=48864;
OC [1]
OC SEQUENCE FROM N.A.
OC MEDLINE=20310618; PubMed=10853771;
RX Slinemann S.J., Anderson O.S., Brown D.W., Miao V.P.;
RA "Cloning and heterologous expression of Solorina crocea PYRG.";
RT Curr. Genet. 37:333-338(2000).
RL -1- CATALYTIC ACTIVITY: Oxidoline 5'-phosphate = UMP + CO(2).
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF206163; AAG34761.1; -.
DR HSSP; P03962; IDOW.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase; 1.
DR PROSITE; PS00156; OMPDECASE; 1.
KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
FT ACT_SITE 95
FT ACT_SITE 95 BY SIMILARITY.
SQ SEQUENCE 280 AA; 30748 MW; 163723FCAB766ACC CRC64;
Query Match 100.0%; Score 22; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOYO 4
DB 226 QOYO 229
RESULT 40
ID PYR_PABR STANDARD; PRT; 285 AA.
AC Q9C31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
 DE (OMPDCase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP
 DE synthase).
 GN URA3.
 OS Paracoccidioides brasiliensis.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Onygenales; Mitosporic Onygenales; Paracoccidioides.
 OX NCBI_TaxID=121759;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IVIC Pb73 / ATCC 32071;
 RA Reinoso C., FERNIGAN E.;
 RT "Isolation and cloning of the URA3 gene of Paracoccidioides
 RT brasiliensis.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate = UMP + CO(2).
 CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
 CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ133782; CAC34740.1; -.
 DR HSRP; P03962; IDQM.
 DR InterPro; IPR001754; OMPdecase.
 DR Pfam; PF00215; OMPdecase; 1.
 DR PROSITE; PS00156; OMPDECASE; 1.
 KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
 FT ACT SITE 95 BY SIMILARITY.
 SQ SEQUENCE 285 AA; 31064 MW; 5B0951F325883DCE CRC64;

Query Match 100.0%; Score 22; DB 1; Length 285;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQYQ 4
 ||||
 Db 233 QQYQ 236

Search completed: August 20, 2003, 09:30:47
 Time : 3.30928 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:19:04 ; Search time 11.6289 Seconds
(without alignments)
88.763 Million cell updates/sec

Title: US-09-991-433-8
Perfect score: 22
Sequence: 1 QOYQ 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_plant:*
- 10: sp_protein:*
- 11: sp_ricinus:*
- 12: sp_vertebrate:*
- 13: sp_virus:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	28	10	Q65311 arabidopsis
2	22	100.0	35	11	Q8BK89 mus musculus
3	22	100.0	41	10	Q96414 daucus caro
4	22	100.0	43	10	Q9M5U2 nicotiana t
5	22	100.0	44	10	Q03996 nicotiana t
6	22	100.0	45	10	Q9MSU1 nicotiana t
7	22	100.0	48	11	Q8VIR3 ratius norv
8	22	100.0	50	16	Q8VKK7 mycobacteri
9	22	100.0	51	5	Q26905 trichostem
10	22	100.0	51	13	Q98TU8 platichthys
11	22	100.0	52	10	Q93XV9 malus domes
12	22	100.0	53	10	Q03995 nicotiana t
13	22	100.0	53	10	Q03997 nicotiana t
14	22	100.0	55	5	Q9G292 plasmodium
15	22	100.0	56	5	Q8ISR0 spodoptera
16	22	100.0	56	10	Q9ZTV7 arabidopsis

17	22	100.0	57	5	Q9G295 plasmodium
18	22	100.0	57	10	Q9S743 arabidopsis
19	22	100.0	58	5	Q9G298 plasmodium
20	22	100.0	58	5	Q9G2A1 plasmodium
21	22	100.0	58	5	Q9G2A7 plasmodium
22	22	100.0	58	10	Q9ZTV6 arabidopsis
23	22	100.0	59	10	Q9ZTV8 arabidopsis
24	22	100.0	60	10	Q9ZTV5 arabidopsis
25	22	100.0	63	5	Q9VCV2 drosophila
26	22	100.0	68	10	Q41946 arabidopsis
27	22	100.0	69	10	Q8H6M2 gossypium h
28	22	100.0	69	16	Q92E04 listeria in
29	22	100.0	71	10	Q96415 daucus caro
30	22	100.0	71	10	Q8GVZ2 pyrus pyrif
31	22	100.0	76	5	Q9VFP9 drosophila
32	22	100.0	77	10	Q39696 daucus caro
33	22	100.0	77	10	Q39616 chenopodium
34	22	100.0	77	12	Q98549 parametium
35	22	100.0	77	16	Q8DLE9 synechococc
36	22	100.0	79	13	Q8UUG0 morone saxa
37	22	100.0	79	13	Q8UUG2 morone chry
38	22	100.0	79	16	Q8FG79 escherichia
39	22	100.0	80	16	Q8Z613 salmonella
40	22	100.0	80	16	Q84943 salmonella
41	22	100.0	80	16	Q8CLX9 druceilla su
42	22	100.0	84	10	Q8H6M9 gossypium h
43	22	100.0	86	16	Q8XTW2 ralsconia s
44	22	100.0	88	10	Q39412 brassica ca
45	22	100.0	89	2	Q9AED7 legionella

ALIGNMENTS

RESULT 1					
ID	Q65311	PRELIMINARY;	PRT;	28	AA.
AC	Q65311;				
DT	01-AUG-1998 (TREMBLrel. 07, Created)				
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	Beta-9 tubulin (Fragment).				
GN	TUB9.				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosis II; Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Ms.				
RA	Sanders P.M., But A.Q., Weterings K., McIntire K.N., Hsu Y.C.,				
RA	Lee P.Y., Truong M.T., Beals T.B., Goldberg R.B.;				
RT	"Anchor Development Defects in Arabidopsis thaliana Male-Sterile				
RT	Mutants."				
RL	Sex. Plant Reprod. 11:297-322(1999).				
DR	EMBL: AF060248; AAC97107.1; .				
FT	NON TER				
SQ	SEQUENCE 28 AA; 3388 MW; 8060E4E537BE670E CRC64;				
Query Match					
Best Local Similarity	100.0%; Score 22; DB 10; Length 28;				
Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 QOYQ 4				
DB	7 QOYQ 10				
RESULT 2					
ID	Q8BK89	PRELIMINARY;	PRT;	35	AA.
AC	Q8BK89;				

DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Tubulin (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxId=10990;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Body;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK05618; BAC35862.1; -.
 FT NON TER 1
 SQ SEQUENCE 35 AA; 4001 MW; 7867F669A738DDB8 CRC64;
 Query Match 100.0%; Score 22; DB 11; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYO 4
 DB 13 QOYO 16
 RESULT 3
 ID 096414 PRELIMINARY; PRT; 41 AA.
 AC 096414;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Beta-tubulin 3 (Fragment).
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
 RN NCBI_TaxId=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Kintoki;
 RA Okamura S., Sonehara K., Naito K., Ohkawa H., Kuramori S., Tatsuta M.,
 RA Minamizono M.;
 RT "Characterization of beta-tubulin genes of carrot."
 RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U64430; AAB47936.1; -.
 FT NON TER 1
 SQ SEQUENCE 41 AA; 4945 MW; A5236AAE5F10274 CRC64;
 Query Match 100.0%; Score 22; DB 10; Length 41;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYO 4
 DB 17 QOYO 20
 RESULT 4
 ID 09MSU2 PRELIMINARY; PRT; 43 AA.
 AC 09MSU2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Beta-tubulin 4 (Fragment).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
 RN NCBI_TaxId=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Okamura S., Hara M., Yamaguchi A.;
 RT "Identification of beta-tubulin isoform gene products in tobacco cells
 and tissues."
 RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF22245; AAF32374.1; -.
 FT NON TER 1
 SQ SEQUENCE 43 AA; 5129 MW; A99AA5E56189119 CRC64;
 Query Match 100.0%; Score 22; DB 10; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYO 4
 DB 21 QOYO 24
 RESULT 5
 ID 003996 PRELIMINARY; PRT; 44 AA.
 AC 003996;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Beta-tubulin 2.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
 RN NCBI_TaxId=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bright Yellow;
 RX MEDLINE=99389269; PubMed=10462173;
 RA Okamura S., Okahara K., Iida T., Ozaki M., Asano S., Morita M.,
 RA Imanaka T.;
 RT "Isotype-specific changes in the amount of beta-tubulin RNA in
 RT synchronized tobacco BY2 cells."
 RT Cell Struct. Funct. 24:117-122(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bright Yellow;
 RA Okamura S., Okahara K., Iida T.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U91563; AAB50566.2; -.
 SQ SEQUENCE 44 AA; 5251 MW; 8840E761EB2AD90 CRC64;
 Query Match 100.0%; Score 22; DB 10; Length 44;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOYO 4
 DB 21 QOYO 24
 RESULT 6
 ID 09MSU1 PRELIMINARY; PRT; 45 AA.
 AC 09MSU1;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Beta-tubulin 5 (Fragment).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.

OK NCBI_TaxId=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okamura S., Hara M., Yamaguchi A.;
 RT "Identification of beta-tubulin isotype gene products in tobacco cells
 and tissues";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF23246; AAF32375.1; -
 RN NON TER
 SQ SEQUENCE 45 AA; 5345 MW; 866CFBCF761BA61F CRC64;

Query Match 100.0%; Score 22; DB 10; Length 45;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 21 QOYQ 24

RESULT 7

OBVIP3 PRELIMINARY; PRT; 48 AA.
 AC OBVIP3;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Beta-tubulin class Iva isotype (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN NCBI_TaxId=10116;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wiater; TISSUE=Brain;
 RA Arai K.;
 RT "Preparation and characterization of a monoclonal antibody to class II
 beta-tubulin isotype.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB062422; BAB72260.1; -
 RN NON TER
 SQ SEQUENCE 48 AA; 5547 MW; 149F8793331629AE CRC64;

Query Match 100.0%; Score 22; DB 11; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 27 QOYQ 30

RESULT 8

OBVKK7 PRELIMINARY; PRT; 50 AA.
 AC OBVKK7;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hypothetical protein MT0494.
 GN MT0494.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 RN NCBI_TaxId=1773;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Bisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Deterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE006951; AAK44717.1; -
 DR TIGR; MT0494; -
 RN Hypothetical protein.
 SQ SEQUENCE 50 AA; 6040 MW; E9756CC5B6F327C3 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 14 QOYQ 17

RESULT 9

Q26905 PRELIMINARY; PRT; 51 AA.
 AC Q26905;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Beta-tubulin (Fragment).
 GN TCB-1.
 OS Trichostromyplus colubriformis.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 OC Trichostromyloidea; Trichostromyloidea; Trichostromylinae;
 OC Trichostromylinus.
 RN NCBI_TaxId=6319;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=McMaster;
 RA Grant W.N., Nascond L.U.;
 RT "Beta-tubulin gene polymorphism and benzimidazole resistance in
 Trichostromyplus colubriformis";
 RL Int. J. Parasitol. 0:0-0(1995).
 DR EMBL; U39620; AAA80231.1; -
 RN NON TER
 SQ SEQUENCE 51 AA; 5750 MW; 78003D1CD26CE60B CRC64;

Query Match 100.0%; Score 22; DB 5; Length 51;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 26 QOYQ 29

RESULT 10

Q98TUB PRELIMINARY; PRT; 51 AA.
 AC Q98TUB;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Utaive E1B-55kDa-associated protein (Fragment).
 GN E1B-AP5.
 OS Platyichthys flesus (European flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorphi; Pleuronectiformes;
 OC Pleuronectoidae; Pleuronectidae; Platyichthys.
 RN NCBI_TaxId=8260;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Williams T.D., Chipman J.K.;
 RT "A DNA array to monitor the effects of environmental pollution on
 European flounder (Platyichthys flesus)";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ310424; CAC27403.1; -.
 FT NON TER 1 1
 SQ SEQUENCE 51 AA; 6108 MW; 803DB0B6FCBCHC84 CRC64;

Query Match 100.0%; Score 22; DB 13; Length 51;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 ||||
 Db 9 00YQ 12

RESULT 11

093XV9 PRELIMINARY; PRT; 52 AA.

ID 093XV9
 AC 093XV9
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)

DE Sh-RNase (Fragment).

OS Malus domestica (apple) (Malus sylvestris).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC euroside 1; Rosales; Rosaceae; Maloideae; Malus.

OX NCBI_TaxID=3750;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Akane; TISSUE=leaf;

RA Matsumoto S., Hayashi S., Kitahara K., Soejima J.;

"Genomic DNA sequences encoding Malus x domestica Borkh. 'Akane' and

'Delicious' S-RNases.";

RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL: A050635; BAB47144.1; -.

DR InterPro: IPR001568; RNase_T2; 1.

DR Pfam: PF00445; ribonuclease_T2; 1.

DR PROSITE: PS00530; RNASE_T2_1; 1.

FT NON TER 1 1

SQ SEQUENCE 52 AA; 5805 MW; B2F42BF7C9D1062D CRC64;

Query Match 100.0%; Score 22; DB 10; Length 52;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 ||||
 Db 40 00YQ 43

RESULT 12

003995 PRELIMINARY; PRT; 53 AA.

ID 003995
 AC 003995
 DT 01-JUL-1997 (TRENBLREL. 04, Created)

DT 01-JUL-1997 (TRENBLREL. 04, Last sequence update)

DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)

DE Beta-tubulin 1 (Fragment).

OS Nicotiana tabacum (Common tobacco).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.

OX NCBI_TaxID=4097;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Bright Yellow;

RX MEDLINE=99389269; PubMed=10462173;

RA Okamura S., Okahara K., Iida T., Ozaki M., Asano S., Morita M.,

Imanaka T.;

"Isotype-specific changes in the amount of beta-tubulin RNA in

synchronized tobacco BY2 cells.";

RL Cell Struct. Funct. 24:117-122(1999).

DR EMBL: U91562; AAB50565.1; -.
 FT NON TER 1 1
 SQ SEQUENCE 53 AA; 6250 MW; FE94091BD7F9DF1B CRC64;

Query Match 100.0%; Score 22; DB 10; Length 53;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 ||||
 Db 29 00YQ 32

RESULT 13

003997 PRELIMINARY; PRT; 53 AA.

ID 003997
 AC 003997
 DT 01-JUL-1997 (TRENBLREL. 04, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)

DE Beta-tubulin 3 (Fragment).

OS Nicotiana tabacum (Common tobacco).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.

OX NCBI_TaxID=4097;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Bright Yellow;

RX MEDLINE=99389269; PubMed=10462173;

RA Okamura S., Okahara K., Iida T., Ozaki M., Asano S., Morita M.,

Imanaka T.;

"Isotype-specific changes in the amount of beta-tubulin RNA in

synchronized tobacco BY2 cells.";

RL Cell Struct. Funct. 24:117-122(1999).

RL [2]

RN SEQUENCE FROM N.A.

RC STRAIN=cv. Bright Yellow;

RA Okamura S., Okahara K., Iida T.;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: U91564; AAB63381.2; -.

FT NON TER 1 1

SQ SEQUENCE 53 AA; 6304 MW; 6977613FD76C9F1B CRC64;

Query Match 100.0%; Score 22; DB 10; Length 53;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 ||||
 Db 29 00YQ 32

RESULT 14

09GZ92 PRELIMINARY; PRT; 55 AA.

ID 09GZ92
 AC 09GZ92
 DT 01-MAR-2001 (TRENBLREL. 16, Created)

DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)

DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)

DE S-antigen (Fragment).

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=T116;

RA Anderson T.J.C., Day K.P.;

"Geographical structure and sequence evolution as inferred from the

Plasmodium falciparum S-antigen locus.";

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF180246; AAC01944.1; -.

FT NON TER 1 1

FT NON_TER 55 55
SQ SEQUENCE 55 AA; 6775 MW; BE8F57176E8797EF CRC64;

Query Match 100.0%; Score 22; DB 5; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
DB 49 QOYQ 52

RESULT 15

ID Q9ISRO PRELIMINARY; PRT; 56 AA.

AC Q9ISRO;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
Beta-1 tubulin (Fragment).
OC Spodoptera frugiperda (Fall armyworm).
OC Buteo borealis; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Nucleoside; Amphipyrinae; Spodoptera.
NCBI_TaxID=7108;
RN [1]
RP SEQUENCE FROM N.A.

RA Noblison I., O'Reilly D.R., Olszewski J.A.;
RT "Differential expression of host cellular genes in Spodoptera frugiperda cells infected with Autographa californica";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF548017; AAN38750.1; -.

FT NON_TER 1 56
SQ SEQUENCE 56 AA; 6764 MW; F181295B16C5753A CRC64;

Query Match 100.0%; Score 22; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
DB 49 QOYQ 52

RESULT 16

ID Q9ZTV7 PRELIMINARY; PRT; 56 AA.

AC Q9ZTV7;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
POSF21 (Fragment).
GN POSF21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-ecotype Bur-0;
RA Lortidon K., Courroyer B., Goubely C., Depelges A., Picard G.;
RT "Length polymorphism of trinucleotide microsatellites in natural accessions of Arabidopsis thaliana";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021623; AAD01860.1; -.

FT NON_TER 1 56
SQ SEQUENCE 56 AA; 7047 MW; 9BF7FB409192816B CRC64;

Query Match 100.0%; Score 22; DB 10; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
DB 37 QOYQ 40

RESULT 17

ID Q9GZ95 PRELIMINARY; PRT; 57 AA.

AC Q9GZ95;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
S-antigen (Fragment).
OC Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.

RA Anderson T.J.C., Day K.P.;
RT "Geographical structure and sequence evolution as inferred from the Plasmodium falciparum S-antigen locus";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF180243; AAG01941.1; -.

FT NON_TER 1 57
SQ SEQUENCE 57 AA; 6919 MW; AFE042E327176E87 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 57;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
DB 49 QOYQ 52

RESULT 18

ID Q9S743 PRELIMINARY; PRT; 57 AA.

AC Q9S743;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
POSF21 (Fragment).
GN POSF21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-ecotype BUR-0;
RA Lortidon K., Courroyer B., Goubely C., Depelges A., Picard G.;
RT "Length polymorphism of trinucleotide microsatellites in natural accessions of Arabidopsis thaliana";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021624; AAD01861.1; -.

FT NON_TER 1 57
SQ SEQUENCE 57 AA; 7175 MW; D78B782C1FC89A43 CRC64;

Query Match 100.0%; Score 22; DB 10; Length 57;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
DB 38 QOYQ 41

RESULT 19

09GZ98 PRELIMINARY; PRT; 58 AA.
 AC 09GZ98;
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
 DE S-antigen (Fragment).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T76;
 RA Anderson T.J.C., Day K.P.;
 RT "Geographical structure and sequence evolution as inferred from the
 RT Plasmodium falciparum S-antigen locus."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF180239; AAC01937.1; --
 FT NON_TER 1 58
 SQ SEQUENCE 58 AA; 7148 MW; B7CD305E8F57176E CRC64;

Query Match 100.0%; Score 22; DB 5; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 49 QOYQ 52

RESULT 20

09GZAI PRELIMINARY; PRT; 58 AA.
 AC 09GZAI;
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
 DE S-antigen (Fragment).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T71;
 RA Anderson T.J.C., Day K.P.;
 RT "Geographical structure and sequence evolution as inferred from the
 RT Plasmodium falciparum S-antigen locus."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF180236; AAC01934.1; --
 FT NON_TER 1 58
 SQ SEQUENCE 58 AA; 6973 MW; 4D9C8322F9971774 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 49 QOYQ 52

RESULT 21

09GZAI PRELIMINARY; PRT; 58 AA.
 AC 09GZAI;
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)

DE S-antigen (Fragment).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=265;
 RA Anderson T.J.C., Day K.P.;
 RT "Geographical structure and sequence evolution as inferred from the
 RT Plasmodium falciparum S-antigen locus."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF180226; AAC01924.1; --
 FT NON_TER 1 58
 SQ SEQUENCE 58 AA; 7091 MW; B89C872F5F5FAA76E CRC64;

Query Match 100.0%; Score 22; DB 5; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 49 QOYQ 52

RESULT 22

09ZTV6 PRELIMINARY; PRT; 58 AA.
 AC 09ZTV6;
 DT 01-MAY-1999 (TEMBLrel. 10, Created)
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
 DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
 DE POSF21 (Fragment).
 GN POSF21.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ecotype Cen-0;
 RA Loidon K., Courtoyer B., Goubely C., Depelges A., Picard G.;
 RT "Length polymorphism of trinuclotide microsatellites in natural
 RT accessions of Arabidopsis thaliana."
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF027625; AAD01862.1; --
 FT NON_TER 1 58
 SQ SEQUENCE 58 AA; 7303 MW; 387704AF6BD751C9 CRC64;

Query Match 100.0%; Score 22; DB 10; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 39 QOYQ 42

RESULT 23

09ZTV8 PRELIMINARY; PRT; 59 AA.
 AC 09ZTV8;
 DT 01-MAY-1999 (TEMBLrel. 10, Created)
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
 DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
 DE POSF21 (Fragment).
 GN POSF21.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ecocyte Ag-0;
 RA Loriclon K., Courmoyer B., Gombely C., Depeiges A., Picard G.;
 RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF027621; AAD01858.1;
 FT NON TER 1 1
 FT 59 59
 SQ SEQUENCE 59 AA; 7431 MW; EE78F8D3B8A34E02 CRC64;
 Query Match 100.0%; Score 22; DB 10; Length 59;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 OQYO 4
 Db 40 OQYO 43
 RESULT 24
 ID 09ZTV5 PRELIMINARY; PRT; 60 AA.
 AC 09ZTV5;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE POSF21. (Fragment).
 GN POSF21.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ecocyte Ka-0;
 RA Loriclon K., Courmoyer B., Gombely C., Depeiges A., Picard G.;
 RL "Length polymorphism of trinucleotide microsatellites in natural
 accessions of Arabidopsis thaliana.";
 RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF027626; AAD01863.1;
 FT NON TER 1 1
 FT 60 60
 SQ SEQUENCE 60 AA; 7559 MW; 55F8F72P94203A1D CRC64;
 Query Match 100.0%; Score 22; DB 10; Length 60;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 OQYO 4
 Db 41 OQYO 44
 RESULT 25
 ID 09VCV2 PRELIMINARY; PRT; 63 AA.
 AC 09VCV2;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE CG17618-PA.
 GN CG17618.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabot G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Plankkoc C., Baldwin D.,
 RA Ballew R.M., Basu A., Bendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Liewicki A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector R., Turner R., Venter B., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassenaar D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodgett M., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banson J., An H., Baldwin D., Banson J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.B., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Mitra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith B., Shu S., Smurnick F., Whitfield B.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF003740; AAF56053.2; -
 DR FlyBase; FBgn0040587; CG17618.
 SQ SEQUENCE 63 AA; 7406 MW; 0950383E3E08EADC CRC64;

Query Match 100.0%; Score 22; DB 5; Length 63;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 ||||
 DB 3 00YQ 6

RESULT 26

Q41946 PRELIMINARY; PRT; 68 AA.

AC Q41946; 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Beta-1 tubulin (Fragment).
 GN Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Seedling;
 RA Desprez T., Amelelem J., Chiappello H., Rouze P., Caboche M., Hoffe H.;
 RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
 EMBL; Z18053; CAA79089.1; -
 FT NON TER 1 1
 SQ SEQUENCE 68 AA; 8106 MW; 5D23FB01CD152D87 CRC64;

Query Match 100.0%; Score 22; DB 10; Length 68;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 ||||
 DB 43 00YQ 46

RESULT 27

Q8H6M2 PRELIMINARY; PRT; 69 AA.

AC Q8H6M2; 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Beta-tubulin 9 (Fragment).
 GN BTUB9.
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Malvales; Malvaceae; Malvoideae; Gossypium.
 OC NCBI_TaxId=3635;

AC [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Xu-142; TISSUE=Fiber;
 RA Ji S., Liang X., Shi Y., Wei G., Lu Y., Zhu Y.;
 RT "High throughput identification, expression, characterization and functional analysis of cotton fiber development-related genes.";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 EMBL; AF521248; AAN32996.1; -

QY 1 00YQ 4
 ||||
 DB 43 00YQ 46

FT NON TER 1 1
 SQ SEQUENCE 69 AA; 8366 MW; BEAB1B6FA039C27E CRC64;

Query Match 100.0%; Score 22; DB 10; Length 69;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 ||||
 DB 48 00YQ 51

RESULT 28

Q92E04 PRELIMINARY; PRT; 69 AA.

AC Q92E04; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein lin0657.
 GN LIN0657.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OC NCBI_TaxId=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIP 11262 / Serovar 6a;
 RX PubMed-11679669;
 RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetoui F., Couve E., de Daruvar A., Denoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Geisel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Krefz J., Kun M., Kuntz F., Kurpkat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,
 RA Nordsted G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RL "Comparative genomics of Listeria species."
 RT Science 294:849-852(2001).
 DR EMBL; AL596165; CAC95889.1; -
 DR ListList; LIN0657; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 69 AA; 8057 MW; 336D52BC28C8BAF28 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 69;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 ||||
 DB 18 00YQ 21

RESULT 29

Q96415 PRELIMINARY; PRT; 71 AA.

AC Q96415; 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Beta-tubulin 4 (Fragment).
 GN Dancus carota (Carrot).
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
 OC NCBI_TaxId=4039;

AC [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Kintoki;
 RA Okamura S., Senohara K., Naito K., Okawa H., Kuramori S., Tatsu M.,
 RA Minamizono M.;

RT "Characterization of beta-tubulin genes of carrot."
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U64431; AAB47937.1; -
 FT NON TER 1
 SQ SEQUENCE 71 AA; 8694 MW; 264680E6A6F6D86 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 22; DB 10; Length 71;
 Pred. No. 2.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 46 QOYQ 49

RESULT 30

08GV22 PRELIMINARY; PRT; 71 AA.

08GV22 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DE S10-RNase (Fragment).

OS Pyrus pyrifolia (Japanese pear) (Pyrus serotina).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC Eucosids I; Rosales; Rosaceae; Maloideae; Pyrus.

NCBI_TaxID=3767;

SEQUENCE FROM N.A.

RC STRAIN=cv. Chensu1r1;

RA Kim H.T., Hirata Y., Nou I.S.;

RT "Discovery of a new S-RNase and determination of S-genotypes by PCR-

RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY158069; AAN76354.1; -

FT NON TER 1

FT NON TER 1

SQ SEQUENCE 71 AA; 8098 MW; 00C9D29A920E3661 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 22; DB 10; Length 71;
 Pred. No. 2.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 6 QOYQ 9

RESULT 31

09VF98 PRELIMINARY; PRT; 76 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE CG14865 protein.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydriidae; Drosophilidae; Drosophila.

NCBI_TaxID=7227;

SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RA Adams M.D., Clinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton R.G., Wortman Q.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heitman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Paclele J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003708; AAF55162.1; -
 FLBase; FBgn0038314; CG14865.

SQ SEQUENCE 76 AA; 8950 MW; 40CCDB4C8BEC509D CRC64;

Query Match
 Best Local Similarity 100.0%; Score 22; DB 5; Length 76;
 Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 10 QOYQ 13

RESULT 32

Q39696 PRELIMINARY; PRT; 77 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE Tubulin like protein (Carrot).

OS Daucus carota (Carrot).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.

NCBI_TaxID=4039;

SEQUENCE FROM N.A.

RA Kaldemoff R., Aleich F., Richter G.;

RL Submitted (JUN-1989) to the EMBL/GenBank/DBJ databases.

DR EMBL; X16608; CAA34609.1; -

FT NON TER 1

FT NON TER 1

SQ SEQUENCE 77 AA; 9371 MW; 686C9E1C1672E0D14 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 22; DB 10; Length 77;
 Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYQ 4
 ||||
 Db 52 QOYQ 55

RESULT 33

039616 PRELIMINARY; PRT; 77 AA.
 ID 039616
 AC 039616
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Tubulin like protein fragment (AA 1-77) (fragment).
 OS Chenopodium rubrum (Red goosefoot) (P5gweed).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Caryophyllales; Chenopodiaceae; Chenopodium.
 NCBI_TaxID=3560;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Kaidenhot R., Aleith F., Richter G.;
 RL Submitted (JUN-1989) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X16700; CAA34673.1; -.
 FT NON TER 1 1
 SEQID 77 AA; 9371 MW; 686C9EC1672E0D14 CRC64;
 Query Match 100.0%; Score 22; DB 10; Length 77;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 Db 52 00YQ 55

RESULT 34

098549 PRELIMINARY; PRT; 77 AA.
 ID 098549
 AC 098549
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE A499L protein.
 GN A499L.
 OS Parametium buraria chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus..
 NCBI_TaxID=10506;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96400190; PubMed=8806566;
 RA Kutch G.F., Li Y., Lu Z., Furuta M., Rock D.L., Van Etten J.L.;
 RT "Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map
 positions 182 to 258."
 RT Virology 223:303-317(1996).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20013326; PubMed=10544099;
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
 RA Lisee A.D., Nickerson K.W., Van Etten J.L.;
 RT "Chlorella virus PBCV-1 encodes a functional homospemidine
 synthase."
 RT Virology 265:254-262(1999).
 RN (3)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20478054; PubMed=11021991;
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
 RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
 PBCV-1." Virology 276:27-36(2000).
 RN (4)
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN (5)
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 DE Van Etten J.L.;

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

[6]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN (7)
 RP SEQUENCE FROM N.A.
 RA Graves M.V., Van Etten J.L.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN (8)
 RP SEQUENCE FROM N.A.
 RA Graves M.V., Van Etten J.L.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN (9)
 RP SEQUENCE FROM N.A.
 RA Gurnon J.R., Graves M.V., Van Etten J.L.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U42580; AAC96866.1; -.
 SEQID 77 AA; 9531 MW; B42882FAF6AD6B19 CRC64;

Query Match 100.0%; Score 22; DB 12; Length 77;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 Db 68 00YQ 71

RESULT 35

08DLB9 PRELIMINARY; PRT; 77 AA.
 ID 08DLB9
 AC 08DLB9
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Tsl0546 protein.
 GN Tsl0546.
 OS Thermosynechococcus elongatus (Thermosynechococcus elongatus).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 NCBI_TaxID=32046;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN=BP-1;
 RC MEDLINE=22225144; PubMed=12240834;
 RX Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
 RA Watanabe A., Itiguchi M., Kawashima K., Kimura T., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
 RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the thermophilic cyanobacterium
 Thermosynechococcus elongatus BP-1."
 RT DNA Res. 9:123-130(2002).
 RL DNA Res. 9:123-130(2002).
 DR EMBL; AP005370; BAC08098.1; -.
 KW Complete proteome.
 SQ SEQUENCE 77 AA; 8888 MW; E85F7A75883A98EB CRC64;

Query Match 100.0%; Score 22; DB 16; Length 77;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 00YQ 4
 Db 38 00YQ 41

RESULT 36

08UUG0 PRELIMINARY; PRT; 79 AA.
 ID 08UUG0
 AC 08UUG0
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Moronecidin.

OC Morone saxatilis (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OC NCBI_TaxID=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lauch X., Shike H., Burns J.C., Westerman M.E., Ostland V.E.,
 RA Carlsberg J.M., Van Olst J.C., Nizet V., Taylor S.W., Shimizu C.,
 RA Bulet P.;
 RT "Discovery and characterization of two isoforms of moronecidin, a
 RT novel antimicrobial peptide from hybrid striped bass.";
 RL J. Biol. Chem. 0:0-0(2002).
 DR EMBL; AF394244; AAL57319.1; -;
 DR EMBL; AF385583; AAL49496.1; -;
 SQ SEQUENCE 79 AA; 9222 MW; CQDFI3ELDAFA7EED CRC64;

Query Match 100.0%; Score 22; DB 13; Length 79;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 ||||
 DB 54 QOYO 57

RESULT 37

OC8UG2 PRELIMINARY; PRT; 79 AA.
 AC Q8UG2;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Moronecidin prepropeptide precursor.
 OS Morone chrysops (White bass).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OC NCBI_TaxID=46259;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lauch X.C., Shike H., Burns J.C., Westerman M.E., Ostland V.E.,
 RA Carlsberg J.M., Van Olst J.C., Nizet V., Taylor S.W., Shimizu C.,
 RA Bulet P.;
 RT "Discovery and characterization of two isoforms of moronecidin, a
 RT novel antimicrobial peptide from hybrid striped bass.";
 RL J. Biol. Chem. 0:0-0(2001).
 DR EMBL; AF394243; AAL57318.1; -;
 DR EMBL; AF394243; AAL57318.1; -;
 KW Signal.
 FT SIGNAL.
 FT CHAIN 23 44 POTENTIAL.
 SQ SEQUENCE 79 AA; 9106 MW; BEF32481C4200EBB CRC64;

Query Match 100.0%; Score 22; DB 13; Length 79;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 ||||
 DB 54 QOYO 57

RESULT 38

OC8FG79 PRELIMINARY; PRT; 79 AA.
 AC Q8FG79;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN C2506.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Raske D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL; AB016762; AAN80962.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 79 AA; 9508 MW; FE56BD7495D02359 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 79;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 ||||
 DB 8 QOYO 11

RESULT 39

OC8Z6L3 PRELIMINARY; PRT; 80 AA.
 AC Q8Z6L3;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Putative secretion system protein.
 GN SSAR OR STY1124.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OC NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.F.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies K.M., Dowd L., White N., Farrar J.,
 RA Ktogh A., Larsen T.S., Leach S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 DR EMBL; AL627271; CAD01969.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 80 AA; 9394 MW; C2FC78B284D4678 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 ||||
 Db 43 QOYO 46

RESULT 40

ID 084943 PRELIMINARY; PRT; 80 AA.
 AC 084943;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative secretory apparatus component (Secretion system effector).
 GN SSAB OR STM1396.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1] SEQUENCE FROM N.A.
 STRAIN=SL1344;
 CITILLO D.W., Valdivia R.H., Monack D., Falkow S.;
 "Macrophage-dependent induction of the Salmonella pathogenicity island 2 type III secretion system and its role in intracellular survival.";
 Mol. Microbiol. 0:0-0(1998).
 RL [2]
 RN [2] SEQUENCE FROM N.A.
 STRAIN=LT2;
 RX MEDLINE=99000132; PubMed=9786193;
 RA Henzel M., Shea J.E., Waterman R., Mundy R., Nikolaus T., Banks G.,
 RA Vazquez-Torres A., Gleeson C., Fang P.C., Holden D.W.;
 "Genes encoding putative effector proteins of the type III secretion system of Salmonella pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages.";
 Mol. Microbiol. 30:163-174(1998).
 RL [3]
 RN [3] SEQUENCE FROM N.A.
 RP STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677669;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.";
 Nature 413:852-856(2001).
 RL EMBL; AF020808; AAC28877.1; -;
 EMBL; AJ224892; CAI2183.1; -;
 EMBL; AE008761; AAL20320.1; -;
 KW Complete proteome.
 SQ SEQUENCE 80 AA; 9366 MW; C2FC62414A04678 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYO 4
 ||||
 Db 43 QOYO 46

Search completed: August 20, 2003, 09:24:03
 Job time: 14.6289 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 68.0412 Seconds
(without alignments)
46.656 Million cell updates/sec

Title: US-09-991-433-48

Perfect score: 110

Sequence: 1 TYFPNKGTQYTDIERPLM 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	370	12	AA13406
2	110	100.0	543	12	AA13405
3	110	100.0	554	16	AAW08987
4	110	100.0	554	20	AAW23230
5	110	100.0	554	24	ABP57264
6	110	100.0	554	24	ABP57267
7	110	100.0	756	21	AAW71231
8	110	100.0	781	16	AAW08986
9	110	100.0	781	20	AAW23227

10	110	100.0	781	24	ABP57263	Human parvovirus B
11	110	100.0	781	24	ABP57266	Human parvovirus B
12	45	40.9	383	22	AA683008	S. epidermidis ope
13	45	40.9	383	22	ABP38856	Staphylococcus epi
14	45	40.9	485	23	AA624847	Infectious hematop
15	45	40.9	767	22	ABG19946	Novel human diagno
16	44	40.0	869	24	ABU66567	Human immunodefici
17	44	40.0	904	22	ABG13993	Human human diagno
18	43.5	39.5	50	22	ABG48210	Human liver peptid
19	43.5	39.5	50	22	ABB28188	Human peptide #839
20	43.5	39.5	50	22	ABB33363	Peptide #869 encod
21	43.5	39.5	50	22	ABB18822	Protein #821 encod
22	43.5	39.5	50	22	AAW54148	Human brain expres
23	43.5	39.5	50	22	AAW65442	Human bone marrow
24	43.5	39.5	50	22	AAW14415	Peptide #849 encod
25	43.5	39.5	50	22	AAW26828	Peptide #865 encod
26	43.5	39.5	50	22	AAW02142	Human peptide encod
27	43.5	39.5	50	22	ABG36194	Human TRAF-2 kinas
28	43.5	39.5	763	19	AAW41734	Human MNB, homolog
29	43.5	39.5	763	22	AAU02221	Rat Dyrk, a homolo
30	43.5	39.5	763	22	AAU02222	Pseudomonas glutar
31	43	39.1	522	23	AAO15899	Mutant Pseudomonas
32	43	39.1	522	23	AAO15900	Mutant Pseudomonas
33	43	39.1	522	23	AAO15901	Mutant Pseudomonas
34	43	39.1	522	23	AAO15902	Mutant Pseudomonas
35	43	39.1	522	23	AAO15903	Mutant Pseudomonas
36	43	39.1	522	23	AAO15904	Mutant Pseudomonas
37	43	39.1	522	23	AAO15905	Mutant Pseudomonas
38	43	39.1	522	23	AAO15906	Mutant Pseudomonas
39	43	39.1	522	23	AAO15907	Mutant Pseudomonas
40	43	39.1	522	23	AAO15908	Mutant Pseudomonas
41	43	39.1	522	23	AAO15909	Mutant Pseudomonas
42	43	39.1	522	23	AAO15910	Mutant Pseudomonas
43	43	39.1	522	23	AAO15911	Mutant Pseudomonas
44	43	39.1	522	23	AAO15912	Mutant Pseudomonas
45	43	39.1	522	23	AAO15913	Mutant Pseudomonas

ALIGNMENTS

RESULT 1	AA13406	standard; Protein; 370 AA.
ID	AA13406	
AC	AA13406;	
XX		
DT	24-OCT-1991 (first entry)	
XX		
DE	Parvo virus B19 PANSE.	
XX		
KW	Primer; PCR; PAPST; globulin.	
OS	Synthetic.	
XX		
PN	DE4003826-A.	
XX		
PD	14-AUG-1991.	
XX		
PF	08-FEB-1990; 90DE-4003826.	
XX		
PR	08-FEB-1990; 90DE-4003826.	
XX		
PA	(MIKRO-) MIKROGEN MOLEKULARB.	
XX		
PI	Schuttschek E, Moltz M;	
XX		
DR	WPI; 1991-246423/34.	
XX		
PT	Immunologically active parvo virus B19 peptide(s) - comprising	
XX	capsid protein VP1 or VP2 fragments, useful for antibody	
XX	detection or vaccination	

PS Claim 16; Page 10-11; 22pp; German.
 XX
 CC VP2 (AAR13405) and its fragments PANSE and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin prepns.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AAR13159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 CC
 XX
 SQ Sequence 370 AA;

Query Match 100.0%; Score 110; DB 12; Length 370;
 Best Local Similarity 100.0%; Pred. No. 6, 7e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIERPLM 20
 |||||
 227 TYFPNKGTOQYTDQIERPLM 246

RESULT 2
 AAR13405
 ID AAR13405 standard; Protein; 543 AA.
 XX
 AC AAR13405;
 XX
 DT 24-OCT-1991 (first entry)
 XX
 DE Parvo virus B19 VP2.
 XX
 KM Primer; PCR; globulin; PANSE; PAPST.
 XX
 CS Synthetic.
 XX
 PN DE4003826-A.
 XX
 PD 14-AUG-1991.
 XX
 PF 08-FEB-1990; 90DE-4003826.
 XX
 PR 08-FEB-1990; 90DE-4003826.
 XX
 PA (MIKR-) MIKROGEN MOLEKULARB.
 XX
 PI Soutschek E, Motz M;
 WPI; 1991-246423/34.
 XX
 PT Immunologically active parvo virus B19 peptide(s) - comprising
 PT capsid protein VP1 or VP2 fragments, useful for antibody
 PT detection or vaccination
 XX
 PS Disclosure; Fig 2-6; 22pp; German.
 XX
 CC VP2 and its fragments PANSE (AAR13406) and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin prepns.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AAR13159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 CC
 XX
 SQ Sequence 543 AA;

Query Match 100.0%; Score 110; DB 12; Length 543;
 Best Local Similarity 100.0%; Pred. No. 1, 1e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIERPLM 20
 |||||
 DB 400 TYFPNKGTOQYTDQIERPLM 419

RESULT 3
 AAM08987
 ID AAM08987 standard; Protein; 554 AA.
 XX
 AC AAM08987;
 XX
 DT 27-FEB-1997 (first entry)
 XX
 DE Human parvovirus VP-2 protein.
 XX
 KM Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 KM non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 KM erythblastemia; abortion; universal fetal hydrops; liver disease;
 KM haemorrhagic fever; rheumatism; detection; IgG antibody.
 XX
 OS Human parvovirus.
 XX
 PN JP07147986-A.
 XX
 PD 13-JUN-1995.
 XX
 PF 24-SEP-1992; 92JP-0281017.
 XX
 PR 24-SEP-1992; 92JP-0281017.
 XX
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PA (ELED-) DENKI KAGAKU KOGYO KK.
 XX
 DR WPI: 1995-242756/32.
 N-PSDB; AAR49535.
 XX

Human parvovirus gene coding for a polypeptide - useful for
 PT developing vaccines against parvoviral diseases such as
 PT erythroblastemia, haemorrhagic fever, etc.
 XX
 PS Claim 3; Page 7-9; 38pp; English.

The sequences given in AAM08986 represent the parvovirus structural
 CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.
 CC
 XX
 SQ Sequence 554 AA;

Query Match 100.0%; Score 110; DB 16; Length 554;
 Best Local Similarity 100.0%; Pred. No. 1, 1e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIERPLM 20
 |||||
 DB 391 TYFPNKGTOQYTDQIERPLM 410

RESULT 4
 AAY23230
 ID AAY23230 standard; Protein; 554 AA.
 XX
 AC AAY23230;
 XX
 DT 26-AUG-1999 (first entry)
 XX
 DE Erythrovirus V9 VP2 protein.
 XX
 KM Erythrovirus V9; differential diagnosis; parvovirus; infection;
 KM erythrovirus screening; typing; immunoassay; VP2 protein.

Query Match 100.0%; Score 110; DB 16; Length 554;
 Best Local Similarity 100.0%; Pred. No. 1, 1e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
XX OS Erythrovirus.
XX PN FR2771751-A1.
XX PD 04-JUN-1999.
XX PF 03-DEC-1997; 97FR-0015197.
XX PR 03-DEC-1997; 97FR-0015197.
XX PA (ASST-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
XX PI Auguste V, Garbary CA, Nguyen QT;
XX DR WPI: 1999-349543/30.
XX DR N-PSDB; ABX81586.
XX PT Erythrovirus V9 and its nucleic acid sequences - can be used in the
    diagnosis of its infections
XX Claim 19; Page 57-58; 80pp; French.
XX The present sequence represents an erythrovirus V9 protein.
CC Probes and primers derived from erythrovirus V9 polynucleotide
CC sequences (ABX81580) can be used for differential diagnosis of
CC erythrovirus (parvovirus) infections by a combination of
CC amplification and hybridisation assay. The probes can also be
CC used to assess susceptibility to erythrovirus infection and
CC for erythrovirus screening and typing. The antibodies can be
CC used in immunoassays for diagnosis of erythrovirus V9 infections.
XX
SQ Sequence 554 AA;
Query Match 100.0%; Score 110; DB 20; Length 554;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TYFPNKGQOYTDQIERPLM 20
    |||||
DB 391 TYFPNKGQOYTDQIERPLM 410

RESULT 5
ABP57264
ID ABP57264 standard; Protein; 554 AA.
XX
XX ABP57264;
    22-APR-2003 (first entry)
XX
DE Human parvovirus B19 clone B1-VP2 amino acid sequence SEQ ID NO:29.
XX
KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
XX
OS Human parvovirus B19.
XX
PN WO2003002753-A2.
XX
PD 09-JAN-2003.
XX
PF 28-JUN-2002; 2002WO-US20684.
XX
PR 28-JUN-2001; 2001US-302077P.
XX
PR 19-MAR-2002; 2002US-365956P.
XX
PR 29-MAR-2002; 2002US-369224P.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Pichuanes S, Shyamala V;
XX
DR WPI: 2003-201510/19.
XX
DR N-PSDB; AB259574.
```

```
XX PT Detecting a human parvovirus B19 infection in a biological sample to
    prevent viral transmission, comprises reacting a parvovirus B19 nucleic
    acid with a primer complementary to the 3'-terminal portion of the RNA
    target sequence
XX
XX Example 4; Fig 7B; 148pp; English.
XX
CC The present invention describes a method for detecting a human parvovirus
CC B19 infection in a biological sample. The method comprises reacting the
CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
CC consisting of a first primer containing a complexing sequence
CC sufficiently complementary to the 3'-terminal portion of the RNA target
CC sequence to complex with. Also described: (1) amplifying a target
CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
CC of 47 700 base pair sequences (see AB259549 to AB259569, and AB259604 to
CC AB259629); (3) a polynucleotide comprising either of 2 4678 base pair
CC sequences (see AB259570 and AB259571); (4) an oligonucleotide primer
CC consisting of a promoter region recognised by a DNA-dependent RNA
CC polymerase operably linked to a human parvovirus B19-specific complexing
CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
CC oligonucleotide primer of (4), and instructions for conducting the
CC diagnostic test. The method is useful for detecting parvovirus infection
CC in a biological sample, such as in blood products, to prevent
CC transmission of the virus through blood and plasma derivatives or by
CC close personal contact. AB259549 to AB259634 and ABP57262 to ABP57267
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 554 AA;
Query Match 100.0%; Score 110; DB 24; Length 554;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TYFPNKGQOYTDQIERPLM 20
    |||||
DB 391 TYFPNKGQOYTDQIERPLM 410

RESULT 6
ABP57267
ID ABP57267 standard; Protein; 554 AA.
XX
XX ABP57267;
    22-APR-2003 (first entry)
XX
DE Human parvovirus B19 clone B6-VP2 amino acid sequence SEQ ID NO:35.
XX
KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
XX
OS Human parvovirus B19.
XX
PN WO2003002753-A2.
XX
PD 09-JAN-2003.
XX
PF 28-JUN-2002; 2002WO-US20684.
XX
PR 28-JUN-2001; 2001US-302077P.
XX
PR 19-MAR-2002; 2002US-365956P.
XX
PR 29-MAR-2002; 2002US-369224P.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Pichuanes S, Shyamala V;
XX
DR WPI: 2003-201510/19.
XX
DR N-PSDB; AB259577.
XX
XX Detecting a human parvovirus B19 infection in a biological sample to
```

PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic acid with a primer complementary to the 3'-terminal portion of the RNA target sequence -

PS Example 4; Fig 10B; 148pp; English.

XX The present invention describes a method for detecting a human parvovirus B19 infection in a biological sample. The method comprises reacting the isolated parvovirus B19 nucleic acid with a first oligonucleotide CC consisting of a first primer containing a complexing sequence CC sufficiently complementary to the 3'-terminal portion of the RNA target CC sequence to complex with. Also described: (1) amplifying a target CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one CC of 47 700 base pair sequences (see AB259549 to AB259569, and AB259604 to CC AB259629); (3) a polynucleotide comprising either of 2 4678 base pair CC sequences (see AB259570 and AB259571); (4) an oligonucleotide primer CC consisting of a promoter region recognised by a DNA-dependent RNA CC polymerase operably linked to a human parvovirus B19-specific complexing CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked to an acridinium ester label; and (6) a diagnostic test kit comprising an CC oligonucleotide primer of (4), and instructions for conducting the CC diagnostic test. The method is useful for detecting parvovirus infection CC in a biological sample, such as in blood products, to prevent CC transmission of the virus through blood and plasma derivatives or by CC close personal contact. AB259549 to AB259634 and AB257262 to AB257267 CC represent sequences used in the exemplification of the present invention.

SQ Sequence 554 AA;

Query Match 100.0%; Score 110; DB 24; Length 554;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIERPLM 20
|||||
DB 391 TYFPNKGTOQYTDQIERPLM 410

RESULT 7
AA71231
ID AAV71231 standard; Protein; 756 AA.

XX AAV71231;
AC AAV71231;
XX
DT 08-SBP-2000 (first entry)
XX

PT Capsid protein encoded by AAV2/human parvovirus B19 chimeric vector.

XX Adeno-associated virus; AAV2; human parvovirus B19; chimeric;
KM recombinant parvoviral vector; cellular tropism; cap protein;
KM capsid; gene delivery; gene therapy; VP1; VP2; VP3.

XX Chimeric - Adeno associated virus serotype 2.
OS Chimeric - Human parvovirus B19.
XX
XX MO200028004-A1.
XX
XX PD 18-MAY-2000.
XX
XX PF 10-NOV-1999; 99WO-US26505.
XX
XX PR 10-NOV-1998; 98US-0107840.
XX
XX PR 10-MAR-1999; 99US-0123651.
XX
XX (UNNC-) UNIV NORTH CAROLINA.
XX
XX PI Rabinowitz JE, Samulski RJ, Xiao W;
XX
XX MPI: 2000-376523/32.
XX
XX DR N-PSDB; AAD00833.
XX
XX PT Recombinant parvoviral vectors with altered packaging, tropisms and

PT immunogenic properties, useful in gene therapy protocols -

XX Example 21; Page 142; 153pp; English.

XX The patent discloses modified parvovirus vectors with advantageous CC antigenic properties, packaging capabilities and cellular tropisms. CC These vectors can be used in standard recombinant DNA protocols e.g. gene CC therapy for delivering nucleic acids to cells.

XX The present sequence is a capsid protein encoded by an CC adeno-associated virus serotype 2 (AAV2)/human parvovirus B19 chimeric CC vector. This vector encodes AAV2 VP1 and VP2 capsid proteins CC and human parvovirus B19 VP2 protein. The chimeric vector was CC constructed by replacing the VP3 major cap protein of AAV2 with B19's CC VP2. Recombinant parvovirus comprising the chimeric capsid is useful CC for gene delivery.

SQ Sequence 756 AA;

Query Match 100.0%; Score 110; DB 21; Length 756;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIERPLM 20
|||||
DB 593 TYFPNKGTOQYTDQIERPLM 612

RESULT 8
AAW08986
ID AAW08986 standard; Protein; 781 AA.

XX AAW08986;
AC AAW08986;
XX
DT 27-FEB-1997 (first entry)
XX
XX DE Human parvovirus VP-1 protein.

XX Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
KM non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
KM erythblastemia; abortion; universal fetal hydrops; liver disease;
KM haemorrhagic fever; rheumatism; detection; IgG antibody.

XX Human parvovirus.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 19
FT Misc-difference /label= Gly, Val
FT Misc-difference 61
FT /label= Asn, Asp
FT Misc-difference 220
FT /label= His, Asn
XX
XX JP07147986-A.
XX
XX PD 13-JUN-1995.
XX
XX PF 24-SEP-1992; 92JP-0281017.
XX
XX PR 24-SEP-1992; 92JP-0281017.
XX
XX (DENK-) DENKA SEIKEN KK.
XX
XX (ELED) DENKI KAGAKU KOGYO KK.
XX
XX MPI: 1995-242756/32.
XX
XX DR N-PSDB; AAT49535.
XX
XX Human parvovirus gene coding for a polypeptide - useful for
PT developing vaccines against parvoviral diseases such as
PT erythroblastemia, haemorrhagic fever, etc.
XX
XX Claim 2; Page 5-7; 38pp; English.

XX The sequences given in AAW08986 represent the parvovirus structural

CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.

XX Sequence 781 AA;

Query Match 100.0%; Score 110; DB 16; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.6e-09; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TYFPNKGTOQYTDQIERPLM 20
 |||||
 Db 618 TYFPNKGTOQYTDQIERPLM 637

RESULT 9

AAV23227 standard; Protein; 781 AA.

AAV23227;

26-AUG-1999 (first entry)

Erythrovirus V9 VP1 protein.

Erythrovirus V9; differential diagnosis; parvovirus; infection;

Erythrovirus screening; typing; immunoassay; VPI protein.

Erythrovirus.

FR2771751-A1.

04-JUN-1999.

03-DEC-1997; 97FR-0015197.

03-DEC-1997; 97FR-0015197.

(ASSISTANCE PUBLIQUE HOPITALUX PARIS.

Auguste V. Garbarg CA, Nguyen QT;

WPI; 1999-349543/30.

N-PSDB; AAX81583.

Erythrovirus V9 and its nucleic acid sequences - can be used in the

diagnosis of its infections

Claim 19; Page 50-52; 80pp; French.

The present sequence represents an erythrovirus V9 protein.

Probes and primers derived from erythrovirus V9 polynucleotide

sequences (AAX81580) can be used for differential diagnosis of

erythrovirus (parvovirus) infections by a combination of

amplification and hybridisation assay. The probes can also be

used to assess susceptibility to erythrovirus infection and

for erythrovirus screening and typing. The antibodies can be

used in immunoassays for diagnosis of erythrovirus V9 infections.

Sequence 781 AA;

Query Match 100.0%; Score 110; DB 20; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.6e-09; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TYFPNKGTOQYTDQIERPLM 20
 |||||
 Db 618 TYFPNKGTOQYTDQIERPLM 637

RESULT 10

ABP57263 standard; Protein; 781 AA.

ABP57263;

22-APR-2003 (first entry)

Human parvovirus B19 clone B1-VP1 amino acid sequence SEQ ID NO:27.

Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.

Human parvovirus B19.

WO2003002753-A2.

09-JAN-2003.

28-JUN-2002; 2002WO-US20684.

28-JUN-2001; 2001US-302077P.

19-MAR-2002; 2002US-365856P.

29-MAR-2002; 2002US-369224P.

(CHIR) CHIRON CORP.

Pichuanes S, Shyamala V;

WPI; 2003-201510/19.

N-PSDB; ABZ59573.

Detecting a human parvovirus B19 infection in a biological sample to

prevent viral transmission, comprises reacting a parvovirus B19 nucleic

acid with a primer complementary to the 3'-terminal portion of the RNA

target sequence

Example 4; Fig 6B; 148bp; English.

The present invention describes a method for detecting a human parvovirus

B19 infection in a biological sample. The method comprises reacting the

isolated parvovirus B19 nucleic acid with a first oligonucleotide

consisting of a first primer containing a complexing sequence

sufficiently complementary to the 3'-terminal portion of the RNA target

sequence to complex with. Also described: (1) amplifying a target

parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one

of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to

ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair

sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer

consisting of a promoter region recognised by a DNA-dependent RNA

polymerase operably linked to a human parvovirus B19-specific complexing

sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a

parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked

to an acridinium ester label; and (6) a diagnostic test kit comprising an

oligonucleotide primer of (4), and instructions for conducting the

diagnostic test. The method is useful for detecting parvovirus infection

in a biological sample, such as in blood products, to prevent

transmission of the virus through blood and plasma derivatives or by

close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267

represent sequences used in the exemplification of the present invention.

Sequence 781 AA;

Query Match 100.0%; Score 110; DB 24; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.6e-09; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TYFPNKGTOQYTDQIERPLM 20
 |||||

Db 618 TYFPNKGTOQYTDQIERPLM 637

RESULT 11

ABP57266

ID ABP57266 standard; Protein; 781 AA.
 XX
 AC ABP57266;
 XX
 DT 22-APR-2003 (first entry)
 XX
 DE Human parvovirus B19 clone B6-VPI amino acid sequence SEQ ID NO:33.
 XX
 DE Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
 XX
 KM Human parvovirus B19;
 XX
 OS Human parvovirus B19.
 XX
 PN WO2003002753-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WO-US20684.
 XX
 PR 28-JUN-2001; 2001US-302077P.
 XX 19-MAR-2002; 2002US-365956P.
 XX 29-MAR-2002; 2002US-369224P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Pichuanes S, Shyamala V;
 XX
 DR WPI; 2003-201510/19.
 XX N-PSDB; ABZ59576.
 DR
 PT Detecting a human parvovirus B19 infection in a biological sample to
 XX prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 XX target sequence -
 XX
 PS Example 4; Fig 9B; 148pp; English.
 XX
 CC The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 781 AA;
 XX
 QY Query Match 100.0%; Score 110; DB 24; Length 781;
 XX Best Local Similarity 100.0%; Pred. No. 1.6e-09;
 XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 TYPNKGTOOYTQIIRPLM 20
 618 TYPNKGTOOYTQIIRPLM 637

RESULT 12
 AAG83008
 ID AAG83008 standard; Protein; 383 AA.
 XX

AC AAG83008;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:3110.
 XX
 DE Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX vaccination; endocarditis.
 XX
 KM Staphylococcus epidermidis.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN WO200134809-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000WO-US30782.
 XX
 PR 09-NOV-1999; 99US-0164258.
 XX
 PA (GLAXO) GLAXO GROUP LTD.
 XX
 PI Kimmery WJ;
 XX
 DR WPI; 2001-316495/33.
 XX N-PSDB; AAH53858.
 DR
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 XX useful for vaccinating against infections, e.g. endocarditis -
 XX
 PS Claim 18; Page 820; 2189pp; English.
 XX
 CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce host cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 SQ Sequence 383 AA;
 XX
 QY Query Match 40.9%; Score 45; DB 22; Length 383;
 XX Best Local Similarity 40.0%; Pred. No. 56;
 XX Matches 12; Conservative 1; Mismatches 5; Indels 12; Gaps 1;
 XX
 DB 2 YFPNKGTOO-----YTQIIRPL 19
 212 YGPOKGAQOKMIPKUDSALRYHDKIEREL 241

RESULT 13
 ABP38856
 ID ABP38856 standard; Protein; 383 AA.
 XX
 AC ABP38856;
 XX
 DT 24-JUL-2002 (first entry)
 XX
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3701.
 XX

KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
XX US6380370-B1.
XX
XX 30-APR-2002.
XX
XX 13-AUG-1998; 98US-0134001.
XX
XX 14-AUG-1997; 97US-055779P.
XX 08-NOV-1997; 97US-064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
XX
XX WPI; 2002-381255/41.
XX N-PSDB; ABN91401.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
XX Disclosure; SEQ ID 3701; 267pp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP512 to ABP3760. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
XX Sequence 383 AA;
SQ
Query Match 40.9%; Score 45; DB 23; Length 383;
Best Local Similarity 40.0%; Pred. No. 56;
Matches 12; Conservative 1; Mismatches 5; Indels 12; Gaps 1;
OY 2 YFPNGTQO-----YTDQIERPL 19
DB 212 YGPKGAQQKMTPKLDSALRHVHDKIEREL 241
AAE24847
ID AAE24847 standard; Protein; 485 AA.
XX
XX AAE24847;
XX
XX 30-OCT-2002 (first entry)
XX
XX Infectious hematopoietic necrosis virus-SCS G protein.
XX
XX G protein; M1 protein; infectious haematopoietic necrosis virus; IHNV;
KW structural protein; vaccine; haematopoietic necrosis symptom; virulence;
KW infectious hematopoietic necrosis virus; IHNV.
XX
XX Infectious hematopoietic necrosis virus.
XX
XX Key Location/Qualifiers
FH Misc-difference 209
FT /note= "Encoded by NC3"
FT Misc-difference 228
FT /note= "Encoded by ATN"
XX
XX WO200236618-A1.
XX
XX

PD 10-MAY-2002.
XX
XX 05-NOV-2001; 2001WO-KR01871.
XX
XX 04-NOV-2000; 2000KR-0065384.
XX
XX (RNAR-) RNA INC.
XX
XX Park J, Moon C, Yoon W, Park K, Kim H, Oh T, Kim H, Chung D;
XX Cho W, Park J;
XX
XX WPI; 2002-519231/55.
XX N-PSDB; AAD40506.
XX
XX New G or M1 structural proteins of a Korean isolate infectious
PT hematopoietic necrosis virus (IHNV-SCS), for use in vaccines against
PT infectious hematopoietic necrosis virus, particularly in fish, e.g.
PT Salmonid fish -
XX
XX Claim 1; Page 54-56; 62pp; English.
XX
XX The present invention relates to novel polynucleotides encoding G or M1
CC structural proteins of infectious haematopoietic necrosis virus (IHNV),
CC which are isolated from Korean type IHNV-SCS. The G or M1 structural
CC proteins of IHNV-SCS are useful as prophylactic vaccines against IHNV.
CC The genes encoding the proteins are useful as DNA vaccines. The vaccines
CC are particularly useful against haematopoietic necrosis symptoms in fish,
CC e.g. Salmonid fish. The present sequence is IHNV-SCS G protein.
XX
XX Sequence 485 AA;
SQ
Query Match 40.9%; Score 45; DB 23; Length 485;
Best Local Similarity 42.1%; Pred. No. 75;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
OY 2 YFPNGTQOYTDQIERPLM 20
DB 144 YFPASRCQWYTDVQNDLI 162
RESULT 15
ABG19946
ID ABG19946 standard; Protein; 767 AA.
XX
XX ABG19946;
XX
XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #19937.
XX
XX Human; chromosome mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS84133.
XX
XX New isolated polynucleotide and encoded polypeptide, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
XX
XX

PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX
 PS Claim 20; SEQ ID No 50305; 103bp; English.
 CC
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 767 AA;

Query Match 40.9%; Score 45; DB 22; Length 767;
 Best Local Similarity 53.8%; Pred. No. 1.3e+02;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 FPNKGQOYTDQI 15
 ||:|||||:
 DB 511 FPNKGTYADQM 523

RESULT 16
 ABU66567
 ID ABU66567 standard; Protein; 869 AA.

AC ABU66567;
 XX
 DT 22-MAY-2003 (first entry)
 XX

DE Human immunodeficiency virus (HIV) envelope (env) protein #3.

XX Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
 XX vaccine; gene therapy; packaging cell line; humoral immune response;
 XX cellular immune response; gene delivery vector; DNA immunisation;
 XX envelope protein; env.
 KM

OS Human immunodeficiency virus.

PN WO2003004657-A1.

PD 16-JAN-2003.

PF 05-JUL-2002; 2002WO-US21421.

PR 05-JUL-2001; 2001US-303192P.

PR 31-AUG-2001; 2001US-316860P.

PR 16-JAN-2002; 2002US-349728P.

PR 16-JAN-2002; 2002US-349733P.

PR 16-JAN-2002; 2002US-349871P.

PA (CHIR) CHIRON CORP.

PI Zur Megele J, Barnett SW, Lian Y;
 XX WPI, 2003-221602/21.
 DR
 XX

PT New synthetic polynucleotides encoding antigenic HIV type B and/or type
 PT C polypeptides, useful as immunogenic compositions or vaccines for
 PT generating humoral or cellular immune responses against HIV in a
 PT subject, especially humans -
 XX
 XX
 PS Disclosure; Fig 2A-C; 262pp; English.

XX The invention describes a synthetic polynucleotide encoding 2 or more
 CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
 CC derived from different HIV subtypes. The polynucleotide is useful for
 CC immunisation, generation of packaging cell lines, or production of HIV
 CC polypeptides. The polynucleotide and its encoded proteins are useful as
 CC immunogenic compositions or vaccines for generating humoral or cellular
 CC immune responses against HIV in a subject, or for inducing neutralising
 CC antibodies against HIV. The gene delivery vector comprising the
 CC polynucleotide is also useful for DNA immunisation of, or for
 CC generating an immune response (e.g. a humoral or cellular immune
 CC response) in, a subject such as a mammal, particularly a human. This
 CC is the amino acid sequence of a human immunodeficiency virus (HIV)
 CC envelope (env) protein.
 CC

SQ Sequence 869 AA;
 Query Match 40.0%; Score 44; DB 24; Length 869;
 Best Local Similarity 36.8%; Pred. No. 2.2e+02;
 Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 TYFPNKGQOYTDQIERPL 19
 ||:|||||:
 DB 399 TYFPNKGTYKXNGNSLPI 417

RESULT 17
 ABG13993
 ID ABG13993 standard; Protein; 904 AA.

AC ABG13993;
 XX

DT 18-FEB-2002 (first entry)
 XX

DE Novel human diagnostic protein #13984.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 KM

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PR (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;
 XX WPI, 2001-639362/73.
 DR N-PSDB; AAS78180.
 XX

DR N-PSDB; AAS78180.

DR N-PSDB; AAS78180.

DR N-PSDB; AAS78180.

DR N-PSDB; AAS78180.

DR N-PSDB; AAS78180.

DR N-PSDB; AAS78180.

DR N-PSDB; AAS78180.

DR N-PSDB; AAS78180.

DR N-PSDB; AAS78180.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX
 PS Claim 20; SEQ ID No 44352; 103bp; English.
 CC
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG3037 represent novel human
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 904 AA;

Query Match

Best Local Similarity: 40.0%; Score 44; DB 22; Length 904;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQI 15
DB 168 TYAPNKGASRFKQV 162

RESULT 18
ABG48210
ID ABG48210 standard; Peptide; 50 AA.

XX ABG48210;

XX 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 26858.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

XX hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analysing gene expression in human adult liver -

XX Claim 27; SEQ ID No 26858; 658bp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for

XX measuring human gene expression in a sample derived from human adult

XX liver, comprising one of 13109 defined nucleotide sequences given in the

CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (I) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABG47348-ABG59930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 50 AA;

SQ Query Match

Best Local Similarity 39.5%; Score 43.5; DB 22; Length 50;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 4 PNKGTOQ-----YTDQIERPL 19
DB 27 PNISDQVSAISYSDIQIQL 47

RESULT 19
ABB28188
ID ABB28188 standard; Peptide; 50 AA.

XX ABB28188;

XX 01-FEB-2002 (first entry)

XX Human peptide #839 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast;

XX disease; cancer.

XX Homo sapiens.

XX WO200157271-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00662.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes,

XX useful for measuring gene expression in sample derived from human

XX breast, comprises number of single exon nucleic acid probes -

XX Claim 27; SEQ ID NO 11156; 327bp + sequence listing; English.

XX The invention relates to a spatially-addressable set of single exon

XX nucleic acid probes for measuring gene expression in a sample derived

XX from human breast and BT 474 cells. The method involves contacting

XX the probes with a collection of detectably labelled nucleic acids

XX derived from mRNA of human breast, and then measuring the label

XX bound to each probe of the microarray. The probes are useful for

XX verifying the expression of regions of genomic DNA predicted to

XX encode proteins. They are useful for gene discovery, and for

XX determining predisposition and/or prognosing breast disease. Gene

expression analysis is useful for assessing the toxicity of chemical agents on cells. The microarray of this invention presents a far greater diversity of probes for measuring gene expression, with far less bias than expressed sequence tag microarrays. The method is suitable for rapid production of functional information from genomic sequence. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 50 AA;

Query Match 39.5%; Score 43.5; DB 22; Length 50;
Best Local Similarity 47.6%; Pred. No. 8.9;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

4 PNRGTQ-----YTDQIERPL 19
27 PNRSDQVSAVSYSQIQPL 47

RESULT 20

ABR33363
ID ABR33363 standard; Peptide; 50 AA.

AC ABR33363;

DT 04-FEB-2002 (first entry)

DE Peptide #869 encoded by human foetal liver single exon probe.

Human; foetal liver; gene expression; single exon nucleic acid probe.

Homo sapiens.

WO200157277-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00669.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human foetal liver -

Claim 27; SEQ ID NO 25998; 639pp + sequence listing; English.

The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 50 AA;

Query Match 39.5%; Score 43.5; DB 22; Length 50;
Best Local Similarity 47.6%; Pred. No. 8.9;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

4 PNRGTQ-----YTDQIERPL 19
27 PNRSDQVSAVSYSQIQPL 47

RESULT 21

ABR18822
ID ABR18822 standard; Protein; 50 AA.

AC ABR18822;

DT 23-JAN-2002 (first entry)

DE Protein #821 encoded by probe for measuring heart cell gene expression.

Human; gene expression; heart; microarray; vascular system; cardiovascular disease; hypertension; cardiac arrhythmia;

congenital heart disease.

Homo sapiens.

WO200157274-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00666.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488899/53.

Single exon nucleic acid probes for analyzing gene expression in human hearts -

Claim 15; SEQ ID NO 20592; 530pp; English.

The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart (see ABR1535-ABR41305). The present sequence is a protein encoded by one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 50 AA;

Query Match 39.5%; Score 43.5; DB 22; Length 50;
Best Local Similarity 47.6%; Pred. No. 8.9;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

4 PNRGTQ-----YTDQIERPL 19
27 PNRSDQVSAVSYSQIQPL 47

```

RESULT 22
AAM54148
ID AAM54148 standard; Protein; 50 AA.
XX
AC AAM54148;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 26253.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
09-AUG-2001.
XX
30-JAN-2001; 2001WO-US00667.
XX
04-FEB-2000; 2000US-0180312.
XX
26-MAY-2000; 2000US-0207456.
XX
30-JUN-2000; 2000US-0608408.
XX
03-AUG-2000; 2000US-0632366.
XX
21-SEP-2000; 2000US-0234687.
XX
27-SEP-2000; 2000US-0236359.
XX
04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-483446/52.
XX
Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
Example 4; SEQ ID NO: 26253; 650pp + Sequence Listing; English.
XX
The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 50 AA;
XX
Query Match 39.5%; Score 43.5; DB 22; Length 50;
Best Local Similarity 47.6%; Pred. No. 8.9;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
OY 4 PNKGTOQ-----YTQIERPL 19
DB 27 PNISDQVSALSYSQIQOPL 47

```

```

XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
DT 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
04-FEB-2000; 2000US-0180312.
XX
26-MAY-2000; 2000US-0207456.
XX
30-JUN-2000; 2000US-0608408.
XX
03-AUG-2000; 2000US-0632366.
XX
21-SEP-2000; 2000US-0234687.
XX
27-SEP-2000; 2000US-0236359.
XX
04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488900/53.
XX
Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
Example 4; SEQ ID NO: 26848; 658pp + Sequence Listing; English.
XX
The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 50 AA;
XX
Query Match 39.5%; Score 43.5; DB 22; Length 50;
Best Local Similarity 47.6%; Pred. No. 8.9;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
OY 4 PNKGTOQ-----YTQIERPL 19
DB 27 PNISDQVSALSYSQIQOPL 47

```

```

RESULT 24
AAM14415
ID AAM14415 standard; Protein; 50 AA.
XX
AC AAM14415;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #849 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
04-FEB-2000; 2000US-0180312.
XX
26-MAY-2000; 2000US-0207456.
XX
30-JUN-2000; 2000US-0608408.
XX
03-AUG-2000; 2000US-0632366.
XX
21-SEP-2000; 2000US-0234687.
XX
27-SEP-2000; 2000US-0236359.

```


PT protein - and related nucleic acid, used to screen for drugs
 PT potentially useful for diagnosis and treatment of immune regulation
 PT disease
 XX
 PS Claim 2, Page 13-15; 21pp; English.
 CC The human tumour necrosis factor receptor associated factor-2 (TRAF2)
 CC kinase protein is one of several TRAF proteins which are involved in the
 CC signalling of various cellular responses including cytotoxicity,
 CC anti-viral activity, immuno-regulatory activities and the
 CC transcriptional regulation of a number of genes. Agents, such as
 CC antibodies, that modulate the binding of TRAF2 to a natural
 CC intracellular binding target are potentially useful in the development
 CC of agents for the diagnosis and treatment of disease involving immune
 CC regulation. Hybridisation probes and primers based on the TRAF2 nucleic
 CC acid sequence can be used diagnostically, e.g. to screen for genetic
 CC mutations or to detect/amplify homologous sequences.
 XX
 SQ Sequence 763 AA;
 QY
 DB Query Match 39.5%; Score 43.5; DB 19; Length 763;
 Best Local Similarity 47.6%; Pred. No. 2.3e+02;
 Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
 4 PNRGTQO-----YTDQIERPL 19
 46 PNRSDQVSAVSYSDDIQQL 66
 RESULT 29
 AAU02221
 ID AAU02221 standard; Protein; 763 AA.
 AC AAU02221;
 DT 23-OCT-2001 (first entry)
 DE Human MNB, homologue of Drosophila minibrain mnb.
 DE Human; MNB; Drosophila melanogaster minibrain; mnb;
 KM Down Syndrome; chromosome 21; mental retardation.
 XX
 OS Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 186 /note= "Encoded by att"
 FT Misc-difference 187 /note= "Encoded by aat"
 FT Misc-difference 188 /note= "Encoded by gcc"
 FT US6251664-BI.
 PN 26-JUN-2001.
 PD 28-JAN-1997; 97US-0789275.
 PF 28-JAN-1997; 97US-0789275.
 PR (PAL/) ESTIVILL PALLEJA X.
 PA
 PI Estivill Palleja X, Pritchard M, Guimera Villaro J;
 PI WPI; 2001-424460/45.
 DR N-PSDB; AAS06346.
 DR
 XX New gene sequence of the Down Syndrome critical region (DSCR) of human
 PT chromosome 21, useful for detecting minibrain expression in DSCR or
 PT studying abnormalities of mental retardation or other defects in
 PT patients with Down syndrome -
 XX
 PS Disclosure; Column 11-16; 16pp; English.

XX
 CC The sequence represents the amino acid sequence of human Down Syndrome
 CC critical region (MNB), a homologue of Drosophila melanogaster minibrain
 CC (mnb). The MNB DNA sequence is useful for detecting minibrain (MNB) in
 CC the Down Syndrome critical region on chromosome 21. The DNA sequence is
 CC particularly useful for determining and studying pathogenic
 CC abnormalities of mental retardation and/or other defects in patients
 CC with Down syndrome.
 XX
 SQ Sequence 763 AA;
 QY
 DB Query Match 39.5%; Score 43.5; DB 22; Length 763;
 Best Local Similarity 47.6%; Pred. No. 2.3e+02;
 Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
 4 PNRGTQO-----YTDQIERPL 19
 46 PNRSDQVSAVSYSDDIQQL 66
 RESULT 30
 AAU02222
 ID AAU02222 standard; Protein; 763 AA.
 AC AAU02222;
 DT 23-OCT-2001 (first entry)
 DE Rat Dyrk, a homologue of Drosophila minibrain mnb.
 DE Rat; MNB; Drosophila melanogaster minibrain; mnb; Dyrk;
 KM Down Syndrome; chromosome 21; mental retardation.
 XX
 OS Rattus sp.
 OS
 PN US6251664-BI.
 PD 26-JUN-2001.
 PF 28-JAN-1997; 97US-0789275.
 PR 28-JAN-1997; 97US-0789275.
 PR (PAL/) ESTIVILL PALLEJA X.
 PA
 PI Estivill Palleja X, Pritchard M, Guimera Villaro J;
 PI WPI; 2001-424460/45.
 DR
 XX New gene sequence of the Down Syndrome critical region (DSCR) of human
 PT chromosome 21, useful for detecting minibrain expression in DSCR or
 PT studying abnormalities of mental retardation or other defects in
 PT patients with Down syndrome -
 XX
 PS Disclosure; Column 15-20; 16pp; English.
 XX
 CC The sequence represents the amino acid sequence of rat minibrain,
 CC Dyrk, a homologue of Drosophila melanogaster minibrain (mnb) and human
 CC Down Syndrome critical region MNB. The Dyrk sequence was used to identify
 CC its human homologue, MNB. The MNB DNA sequence is useful for detecting
 CC minibrain (MNB) in the Down Syndrome critical region on chromosome 21.
 CC The DNA sequence is particularly useful for determining and studying
 CC pathogenic abnormalities of mental retardation and/or other defects in
 CC patients with Down syndrome.
 XX
 SQ Sequence 763 AA;
 QY
 DB Query Match 39.5%; Score 43.5; DB 22; Length 763;
 Best Local Similarity 47.6%; Pred. No. 2.3e+02;
 Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
 4 PNRGTQO-----YTDQIERPL 19
 46 PNRSDQVSAVSYSDDIQQL 66


```

Db          46 FNISDQVSALSYSDIQOPL 66

RESULT 31
AA015899
ID    AA015899 standard; protein; 522 AA.
XX
XX
AC    AA015899;
DT
DT      30-JAN-2003 (first entry)
DE      Pseudomonas glutaryl amidase (GA) enzyme B chain.
XX
XX      Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA;
KM      7-amino-cephalosporanic acid.
XX
OS      Pseudomonas sp.
XX
PN      MO200272806-AA2.
XX
XX      19-SEP-2002.
PF      12-MAR-2002; 2002WO-IB02119.
XX
PR      14-MAR-2001; 2001DE-1012608.
PR      02-OCT-2001; 2001DE-1048723.
PR      31-OCT-2001; 2001DE-1053389.
XX
XX      (PLAC ) MAX PLANCK GBS FOERDERUNG WISSENSCHAFTEN.
PA      (KOLL/) KOLLER K.
PA      (LANG/) LANGE G.
PA      (SAUB/) SAUBER K.
PI      Koller K, Lange G, Sauber K, Fritz-wolf K, Kabach W;
XX
DR      WPI; 2002-732828/79.
XX
PT      New non-naturally occurring variant of glutaryl amidase comprising
PT      histidine or glutamate in its substrate binding pocket, binds
XX      cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
PT      7-aminocephalosporanic acid
XX
PS      Claim 2; Page 28-30; 36pp; English.
CC
CC      The invention comprises the amino acid sequences of mutant Pseudomonas
CC      glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC      bind cephalosporin C (CPC) as a substrate and catalyse the conversion of
CC      CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC      the invention comprise a histidine or glutamate residue in the substrate
CC      binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
CC      mutant GA proteins of the invention are useful for preparing 7-ACA from
CC      CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
CC      present amino acid sequence represents the B chain of the Pseudomonas
CC      glutaryl amidase enzyme.
XX
SQ      Sequence . 522 AA:
Query Match           39.1%; Score 43; DB 23; Length 522;
Best Local Similarity 72.7%; Pred. No. 1.8e+02;
Matches      8; Conservative      1; Mismatches      2; Indels      0; Gaps      0
OY      7 GTGQYTQDIQR 17
        ||| :|||
DB      482 GTTHYSQDIQR 492

RESULT 32
AA015900
ID    AA015900 standard; protein; 522 AA.
XX
XX
AC    AA015900;
DT
DT      30-JAN-2003 (first entry)

```

```

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #1.
XX
XX Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
KM glutaryl-7-amino-cephalosporanic acid.
XX
XX Pseudomonas sp.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 24 /note= "Wild type Leu replaced by Arg"
FT FT
FT Misc-difference 57 /note= "Wild type Arg replaced by His"
FT FT
FT Misc-difference 177 /note= "Wild type Phe replaced by His"
FT FT
XX WO200272806-A2.
XX
XX 19-SEP-2002.
XX
XX 12-MAR-2002; 2002WO-IB02119.
XX
XX 14-MAR-2001; 2001DE-1012608.
XX 02-OCT-2001; 2001DE-1048723.
XX 31-OCT-2001; 2001DE-1053389.
XX
XX (PLAC) MAX PLANCK GES FORBBERUNG WISSENSCHAFTEN.
XX (KOLL) KOLLER K.
XX (LANG/) LANGE G.
XX (SAMB/) SAUBER K.
XX
XX Koller K, Lange G, Sauber K, Fritz-wolf K, Kabsch W,
PI WPI; 2002-732828/79.
XX
XX New non-naturally occurring variant of glutaryl amidase comprising
PT histidine or glutamate in its substrate binding pocket, binds
PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
PT 7-aminocephalosporanic acid -
XX
XX Claim 10; Page -; 36pp; English.
XX
XX The invention comprises the amino acid sequences of mutant Pseudomonas
CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC bind cephalosporin C (CPC) as a substrate and catalyse the conversion of
CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC the invention comprise a histidine or glutamate residue in the substrate
CC binding pocket that binds the alpha-amino adipy moiety of the CPC. The
CC mutant GA proteins of the invention are useful for preparing 7-ACA from
CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
CC present amino acid sequence represents a mutant B chain of the
CC Pseudomonas glutaryl amidase enzyme.
CC
CC NOTE: The present sequence is not shown in the specification but is
CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
CC AAO15899).
XX
XX Sequence 522 AA;
SQ
XX
XX Query Match 39.1%; Score 43; DB 23; Length 522;
XX Best Local Similarity 72.7%; Fred. No. 1.8e+02;
XX Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0.
XX
XX 7 GTGQYTDQIER 17
XX |||:|||||
XX 482 GTTHXSQIER 492
XX
XX RESULT 33
XX AAO15901
XX ID AAO15901 standard; protein; 522 AA.
XX
XX AAO15901;
AC

```

```

XX 30-JAN-2003 (first entry)
XX Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #2.
XX
XX Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
XX muten; 7-amino-cephalosporanic acid.
XX
XX Pseudomonas sp.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 24 /note= "Wild type Leu replaced by Arg"
XX FT Misc-difference 50 /note= "Wild type Gln replaced by Val"
XX FT Misc-difference 57 /note= "Wild type Arg replaced by His"
XX FT Misc-difference 177 /note= "Wild type Phe replaced by His"
XX FT
XX
XX WO200272806-A2.
XX
XX 19-SEP-2002.
XX
XX 12-MAR-2002; 2002WO-IB02119.
XX
XX 14-MAR-2001; 2001DE-1012608.
XX 02-OCT-2001; 2001DE-1048723.
XX 31-OCT-2001; 2001DE-1053389.
XX
XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (KOLL/) KOLLER K.
XX (LANG/) LANGE G.
XX (SAUB/) SAUBER K.
XX
XX Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
XX
XX WPI; 2002-732828/79.
XX
XX New non-naturally occurring variant of glutaryl amidase comprising
XX histidine or glutamate in its substrate binding pocket, binds
XX cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX 7-aminocephalosporanic acid -
XX
XX Claim 10; Page -; 36pp; English.
XX
XX The invention comprises the amino acid sequences of mutant Pseudomonas
XX glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
XX bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
XX CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
XX the invention comprise a histidine or glutamate residue in the substrate
XX binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
XX mutant GA proteins of the invention are useful for preparing 7-ACA from
XX CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
XX present amino acid sequence represents a mutant B chain of the
XX Pseudomonas glutaryl amidase enzyme.
XX NOTE: The present sequence is not shown in the specification but is
XX derived from the Pseudomonas wild type glutaryl amidase B chain, (see
XX AAO15899).
XX
XX Sequence 522 AA;
XX
XX Query Match 39.1%; Score 43; DB 23; Length 522;
XX Best Local Similarity 72.7%; Pred. No. 1.8e+02;
XX Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX 7 GTGQYTDQIER 17
XX |||:|||||
XX Db 482 GTTHVSDQIER 492
XX
XX RESULT 34

```

```

AAO15902
ID AAO15902 standard; protein; 522 AA.
XX
XX AAO15902;
XX
XX 30-JAN-2003 (first entry)
XX
XX Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #3.
XX
XX Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
XX muten; 7-amino-cephalosporanic acid.
XX
XX Pseudomonas sp.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 24 /note= "Wild type Leu replaced by Arg"
XX FT Misc-difference 50 /note= "Wild type Gln replaced by Ala"
XX FT Misc-difference 57 /note= "Wild type Arg replaced by His"
XX FT Misc-difference 177 /note= "Wild type Phe replaced by His"
XX FT
XX
XX WO200272806-A2.
XX
XX 19-SEP-2002.
XX
XX 12-MAR-2002; 2002WO-IB02119.
XX
XX 14-MAR-2001; 2001DE-1012608.
XX 02-OCT-2001; 2001DE-1048723.
XX 31-OCT-2001; 2001DE-1053389.
XX
XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (KOLL/) KOLLER K.
XX (LANG/) LANGE G.
XX (SAUB/) SAUBER K.
XX
XX Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
XX
XX WPI; 2002-732828/79.
XX
XX New non-naturally occurring variant of glutaryl amidase comprising
XX histidine or glutamate in its substrate binding pocket, binds
XX cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX 7-aminocephalosporanic acid -
XX
XX Claim 10; Page -; 36pp; English.
XX
XX The invention comprises the amino acid sequences of mutant Pseudomonas
XX glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
XX bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
XX CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
XX the invention comprise a histidine or glutamate residue in the substrate
XX binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
XX mutant GA proteins of the invention are useful for preparing 7-ACA from
XX CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
XX present amino acid sequence represents a mutant B chain of the
XX Pseudomonas glutaryl amidase enzyme.
XX NOTE: The present sequence is not shown in the specification but is
XX derived from the Pseudomonas wild type glutaryl amidase B chain, (see
XX AAO15899).
XX
XX Sequence 522 AA;
XX
XX Query Match 39.1%; Score 43; DB 23; Length 522;
XX Best Local Similarity 72.7%; Pred. No. 1.8e+02;
XX Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX 7 GTGQYTDQIER 17
XX |||:|||||
XX

```

Db	482 GTTHXSDQIER 492
RESULT 35	
ID	AAO15903 standard; protein; 522 AA.
XX AC	AAO15903;
XX XX	30-JAN-2003 (first entry)
DT	Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #4.
XX DE	
XX XX	Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
KW	murein; 7-amino-cephalosporanic acid.
KM	
XX	Pseudomonas sp.
OS	Synthetic.
XX OS	
Key	Location/Qualifiers
●	Misc-difference 24 /note= "Wild type Leu replaced by Arg"
FT	Misc-difference 50 /note= "Wild type Gln replaced by Val"
FT	Misc-difference 57 /note= "Wild type Arg replaced by His"
FT	Misc-difference 70 /note= "Wild type Val replaced by His"
FT	Misc-difference 177 /note= "Wild type Phe replaced by His"
XX FT	
XX FN	WO200272806-A2.
XX PD	19-SEP-2002.
PF	12-MAR-2002; 2002WO-IB02119.
PR	14-MAR-2001; 2001DE-1012608.
PR	02-OCT-2001; 2001DE-1048723.
PR	31-OCT-2001; 2001DE-1053389.
XX PA	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA PA	(KOLL/) KOLLER K.
PA PA	(LANG/) LANGE G.
PI	(SAUB/) SAUBER K.
PI	Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
WPI; 2002-732828/79.	
PT	New non-naturally occurring variant of glutaryl amidase comprising
PT	histidine or glutamate in its substrate binding pocket, binds
PT	cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX PT	7-aminocephalosporanic acid -
PS	
PS	Claim 10; Page -: 36pp; English.
XX	The invention comprises the amino acid sequences of mutant Pseudomonas
CC	glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC	bind cephalosporin C (CPC) as a substrate and catalyse the conversion of
CC	CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC	the invention comprise a histidine or glutamate residue in the substrate
CC	binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
CC	mutant GA proteins of the invention are useful for preparing 7-ACA from
CC	CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
CC	present amino acid sequence represents a mutant B chain of the
CC	Pseudomonas glutaryl amidase enzyme.
CC	NOTE: The present sequence is not shown in the specification but is
CC	derived from the Pseudomonas wild type glutaryl amidase B chain, (see
CC	AAO15899).
XX	
Sequence	522 AA;

```

Query March 39.1%; Score 43; DB 23; Length 522;
Best Local Similarity 72.7%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 7 GTQOQYTDQIER 17
    |||:|||||
Db 482 GTTHYSDQIER 492

RESULT 36
AAOI5904
ID AAOI5904 standard; protein; 522 AA.
XX
AC AAOI5904;
XX
DT 30-JAN-2003 (first entry)
XX
DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #5.
XX
KM [Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
XX muten; 7-amino-cephalosporanic acid.
OS Pseudomonas sp.
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Misc-difference 24 /note= "Wild type Leu replaced by Arg"
FT /note= "Wild type Gln replaced by Ala"
FT Misc-difference 50 /note= "Wild type Arg replaced by His"
FT /note= "Wild type Val replaced by His"
FT Misc-difference 70 /note= "Wild type Val replaced by His"
FT /note= "Wild type Phe replaced by His"
FT Misc-difference 177
FT /note= "Wild type Phe replaced by His"
XX
PM WO200272806-A2.
XX
PD 19-SEP-2002.
XX
PE 12-MAR-2002; 2002WO-IB02119.
XX
PR 14-MAR-2001; 2001DE-1012608.
PR 02-OCT-2001; 2001DE-1048723.
PR 31-OCT-2001; 2001DE-1053389.
XX
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (KOLL//) KOLLER K.
PA (LANG//) LANGE G.
PA (SAUB//) SAUBER K.
PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabach W;
XX
XX WPI; 2002-732828/79.
XX
XX New non-naturally occurring variant of glutaryl amidase comprising
XX histidine or glutamate in its substrate binding pocket, binds
XX cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX 7-aminocephalosporanic acid -
XX
XX Claim 10; Page -; 36pp; English.
XX
XX The invention comprises the amino acid sequences of mutant Pseudomonas
XX glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
XX bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
XX CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
XX the invention comprise a histidine or glutamate residue in the substrate
XX binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
XX mutant GA proteins of the invention are useful for preparing 7-ACA from
XX CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
XX present amino acid sequence represents a mutant B chain of the
XX pseudomonas glutaryl amidase enzyme.

```

CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AA015899).

SO Sequence 522 AA;

Query Match 39.1%; Score 43; DB 23; Length 522;
 Best Local Similarity 72.7%; Pred. No. 1.8e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQGYTDQIER 17
 |||:|||||
 Db 482 GTTHYSDQIER 492

RESULT 37
 AA015905
 ID AA015905 standard; protein; 522 AA.
 XX AA015905;

30-JAN-2003 (first entry)

XX Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #6.
 DE Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 KM muterin; 7-amino-cephalosporanic acid.
 XX Pseudomonas sp.
 OS Synthetic.

XX Key Location/Qualifiers

FT MISC-difference 24 /note= "Wild type Leu replaced by Arg"
 FT MISC-difference 50 /note= "Wild type Gln replaced by Val"
 FT MISC-difference 57 /note= "Wild type Arg replaced by His"
 FT MISC-difference 70 /note= "Wild type Val replaced by His"
 FT MISC-difference 153 /note= "Wild type Tyr replaced by Leu"
 FT MISC-difference 177 /note= "Wild type Phe replaced by His"
 FT MISC-difference 177 /note= "Wild type Phe replaced by His"
 PN WO200272806-A2.
 XX 19-SEP-2002.
 XX 12-MAR-2002; 2002MO-IB02119.
 XX 14-MAR-2001; 2001DE-1012608.
 PR 02-OCT-2001; 2001DE-1048723.
 PR 31-OCT-2001; 2001DE-1053389.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (KOLL/) KOLLER K.
 PA (LANG/) LANGER G.
 PA (SAUB/) SAUBER K.

PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
 DR WPI; 2002-732828/79.

XX New non-naturally occurring variant of glutaryl amidase comprising
 PT histidine or glutamate in its substrate binding pocket, binds
 PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
 PT 7-aminocephalosporanic acid
 PS Claim 10; Page -; 36pp; English.

XX The invention comprises the amino acid sequences of mutant Pseudomonas
 CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention

CC bind cephalosporin C (CPC) as a substrate and catalyse the conversion of
 CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
 CC the invention comprise a histidine or glutamate residue in the substrate
 CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.
 CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AA015899).

SO Sequence 522 AA;

Query Match 39.1%; Score 43; DB 23; Length 522;
 Best Local Similarity 72.7%; Pred. No. 1.8e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQGYTDQIER 17
 |||:|||||
 Db 482 GTTHYSDQIER 492

RESULT 38
 AA015906
 ID AA015906 standard; protein; 522 AA.
 XX AA015906;

30-JAN-2003 (first entry)

XX Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #7.
 DE Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 KM muterin; 7-amino-cephalosporanic acid.
 XX Pseudomonas sp.
 OS Synthetic.

XX Key Location/Qualifiers

FT MISC-difference 24 /note= "Wild type Leu replaced by Arg"
 FT MISC-difference 50 /note= "Wild type Gln replaced by Ala"
 FT MISC-difference 57 /note= "Wild type Arg replaced by His"
 FT MISC-difference 70 /note= "Wild type Val replaced by His"
 FT MISC-difference 153 /note= "Wild type Tyr replaced by Leu"
 FT MISC-difference 177 /note= "Wild type Phe replaced by His"
 FT MISC-difference 177 /note= "Wild type Phe replaced by His"
 PN WO200272806-A2.

XX 19-SEP-2002.

XX 12-MAR-2002; 2002MO-IB02119.

XX 14-MAR-2001; 2001DE-1012608.
 PR 02-OCT-2001; 2001DE-1048723.
 PR 31-OCT-2001; 2001DE-1053389.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (KOLL/) KOLLER K.
 PA (LANG/) LANGER G.
 PA (SAUB/) SAUBER K.

PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
 DR WPI; 2002-732828/79.

XX New non-naturally occurring variant of glutaryl amidase comprising

PT histidine or glutamate in its substrate binding pocket, binds
 PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
 PT 7-aminocephalosporanic acid -

PS Claim 10; Page -; 36pp; English.

CC The invention comprises the amino acid sequences of mutant Pseudomonas
 CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
 CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
 CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
 CC the invention comprise a histidine or glutamate residue in the substrate
 CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.
 CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AAO15899).

Sequence 522 AA;

Query Match 39.1%; Score 43; DB 23; Length 522;
 Best Local Similarity 72.7%; Pred. No. 1.8e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQGYTDQIER 17
 |||:||||
 Db 482 GTTHYSQDIER 492

RESULT 39

AAO15907
 ID AAO15907 standard; protein; 522 AA.

AC AAO15907;

DT 30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #8.

KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 KW mutein; 7-amino-cephalosporanic acid.

OS Pseudomonas sp.

OS Synthetic.

XX Key Location/Qualifiers

Misc-difference 176 /note= "Wild type Thr replaced by Asp"

FT Misc-difference 177 /note= "Wild type Phe replaced by His"

FN WO200272806-A2.

PD 19-SEP-2002.

PF 12-MAR-2002; 2002WO-1B02119.

PR 14-MAR-2001; 2001DE-1012608.

PR 02-OCT-2001; 2001DE-1048723.

PR 31-OCT-2001; 2001DE-1053389.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (KOLL/) KOLLER K.

PA (LANG/) LANGE G.

PA (SAUB/) SAUBER K.

PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;

DR WPI; 2002-732828/79.

XX New non-naturally occurring variant of glutaryl amidase comprising

PT histidine or glutamate in its substrate binding pocket, binds
 PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
 PT 7-aminocephalosporanic acid -

PS Claim 10; Page -; 36pp; English.

CC The invention comprises the amino acid sequences of mutant Pseudomonas
 CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
 CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
 CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
 CC the invention comprise a histidine or glutamate residue in the substrate
 CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.
 CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AAO15899).

Sequence 522 AA;

Query Match 39.1%; Score 43; DB 23; Length 522;
 Best Local Similarity 72.7%; Pred. No. 1.8e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQGYTDQIER 17
 |||:||||
 Db 482 GTTHYSQDIER 492

RESULT 40

AAO15908
 ID AAO15908 standard; protein; 522 AA.

AC AAO15908;

DT 30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #9.

KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 KW mutein; 7-amino-cephalosporanic acid.

OS Pseudomonas sp.

OS Synthetic.

XX Key Location/Qualifiers

Misc-difference 24 /note= "Wild type Leu replaced by Arg"

FT Misc-difference 57 /note= "Wild type Arg replaced by His"

FT Misc-difference 176 /note= "Wild type Thr replaced by Asp"

FT Misc-difference 177 /note= "Wild type Phe replaced by His"

FN WO200272806-A2.

PD 19-SEP-2002.

PF 12-MAR-2002; 2002WO-1B02119.

PR 14-MAR-2001; 2001DE-1012608.

PR 02-OCT-2001; 2001DE-1048723.

PR 31-OCT-2001; 2001DE-1053389.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (KOLL/) KOLLER K.

PA (LANG/) LANGE G.

PA (SAUB/) SAUBER K.

PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;

XX
DR WPI; 2002-732828/79.

XX
PT New non-naturally occurring variant of glutaryl amidase comprising
PT histidine or glutamate in its substrate binding pocket, binds
PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
PT 7-aminocephalosporanic acid -

XX
PS Claim 10; Page -; 36pp; English.

XX
CC The invention comprises the amino acid sequences of mutant Pseudomonas
CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC the invention comprise a histidine or glutamate residue in the substrate
CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
CC mutant GA proteins of the invention are useful for preparing 7-ACA from
CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
CC present amino acid sequence represents a mutant B chain of the
CC Pseudomonas glutaryl amidase enzyme.

CC NOTE: The present sequence is not shown in the specification but is
CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
CC AA015899).

XX
SQ Sequence 522 AA;

Query Match 39.1%; Score 43; DB 23; Length 522;

Best Local Similarity 72.7%; Pred. No. 1.8e+02;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQQTDTQIER 17

DB 482 GTTHYSQIER 492

Search completed: August 20, 2003, 09:29:42
Job time : 69.0412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:19:07 ; Search time 22.0619 Seconds
(without alignments)
38.357 Million cell updates/sec

Title: US-09-991-433-48

Perfect score: 110

Sequence: 1 TYPNKGTYQYTDIERPLM 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	395	US-08-856-841-13	Sequence 13, Appl
2	110	100.0	398	US-08-856-841-21	Sequence 21, Appl
3	110	100.0	415	US-08-856-841-20	Sequence 20, Appl
4	110	100.0	543	US-08-856-841-22	Sequence 22, Appl
5	110	100.0	756	US-09-438-268-4	Sequence 4, Appl
6	45	40.9	383	US-09-134-001C-3701	Sequence 3701, Ap
7	44	40.0	119	US-09-345-236B-135	Sequence 135, App
8	44	40.0	235	US-09-345-236B-108	Sequence 108, App
9	43.5	39.5	763	US-08-677-862-2	Sequence 2, Appl
10	43.5	39.5	763	US-09-252-571-2	Sequence 2, Appl
11	43.5	39.5	763	US-09-434-065-2	Sequence 2, Appl
12	43.5	39.5	763	US-08-789-275-4	Sequence 4, Appl
13	43.5	39.5	763	US-08-789-275-5	Sequence 5, Appl
14	43	39.1	720	US-07-731-157A-2	Sequence 2, Appl
15	43	39.1	720	US-08-541-780-2	Sequence 2, Appl
16	42	38.2	655	US-09-252-991A-17112	Sequence 17112, A
17	42	38.2	1501	US-08-447-464-3	Sequence 3, Appl
18	42	38.2	1501	US-08-716-679-3	Sequence 3, Appl
19	42	38.2	1911	US-08-348-006B-5	Sequence 5, Appl
20	42	38.2	1911	US-08-800-825A-5	Sequence 5, Appl
21	42	38.2	1911	US-09-158-657-5	Sequence 5, Appl
22	42	38.2	1911	PCT-US94-1016-5	Sequence 5, Appl
23	40	36.4	187	US-09-095-855-203	Sequence 203, App
24	40	36.4	187	US-09-205-426-203	Sequence 205, App
25	40	36.4	1097	US-09-252-991A-22579	Sequence 22579, A
26	40	36.4	1257	US-09-220-641-3	Sequence 3, Appl
27	40	36.4	1872	US-08-188-582-14	Sequence 14, Appl

28	40	36.4	1872	1	US-08-646-715-14	Sequence 14, Appl
29	40	36.4	1893	1	US-08-188-582-11	Sequence 11, Appl
30	40	36.4	1893	1	US-08-646-715-11	Sequence 11, Appl
31	39.5	35.9	151	4	US-09-732-210-1674	Sequence 1674, Ap
32	39.5	35.9	153	4	US-09-732-210-1672	Sequence 1672, Ap
33	39.5	35.9	510	1	US-08-249-112-3	Sequence 3, Appl
34	39.5	35.9	510	1	PCT-US95-06556-3	Sequence 3, Appl
35	39	35.5	16	3	US-08-931-220-2	Sequence 2, Appl
36	39	35.5	16	5	PCT-US95-11723-2	Sequence 2, Appl
37	39	35.5	16	5	PCT-US96-05897-15	Sequence 15, Appl
38	39	35.5	124	4	US-09-107-532A-4773	Sequence 4773, Ap
39	39	35.5	166	4	US-08-311-731A-299	Sequence 299, App
40	39	35.5	241	4	US-09-134-001C-5493	Sequence 5493, Ap
41	39	35.5	322	4	US-09-326-806-3	Sequence 3, Appl
42	39	35.5	398	3	US-08-931-220-5	Sequence 5, Appl
43	39	35.5	398	5	PCT-US96-11723-5	Sequence 5, Appl
44	39	35.5	398	5	PCT-US96-05897-1	Sequence 1, Appl
45	39	35.5	477	4	US-09-328-352-5428	Sequence 5428, Ap

ALIGNMENTS

RESULT 1
US-08-856-841-13
; Sequence 13, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HADIT HAFNER & DELAUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856, 841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214, 658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US/07/917, 096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE

```

: HYPOTHEICAL: N/A
: ANTI-SENSE: N/A
: FRAGMENT TYPE: INTERNAL
: ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
: ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
: IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
: POSITION IN GENOME: N/A
: FEATURE:
: NAME/KEY: N/A
: LOCATION: N/A
: IDENTIFICATION METHOD: amino acid analysis and
: IDENTIFICATION METHOD: mass spectrometry
: OTHER INFORMATION:
: PUBLICATION INFORMATION:
: AUTHORS: COSSART, Y.E.
: AUTHORS: FIELD, A.M.
: AUTHORS: CANT, B.
: AUTHORS: WIDDOVS, D.
: TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
: JOURNAL: LANCET
: VOLUME: 1
: ISSUE:
: PAGES: 72 - 73
: DATE: 1975
: DOCUMENT NUMBER:
: FILING DATE:
: PUBLICATION DATE:
: RELEVANT RESIDUES IN SEQ ID NO: 13:
: PUBLICATION INFORMATION:
: AUTHORS: MANIATIS, T.
: AUTHORS: PRITSCH, E.F.
: AUTHORS: SAMBROOK, J.
: TITLE: MOLECULAR CLONING
: JOURNAL: COLD SPRING HARBOR, NY
: VOLUME:
: ISSUE:
: PAGES:
: DATE: 1982
: DOCUMENT NUMBER:
: FILING DATE:
: PUBLICATION DATE:
: RELEVANT RESIDUES IN SEQ ID NO: 13:
: PUBLICATION INFORMATION:
: AUTHORS: SMITH, D.B.
: AUTHORS: JOHNSON, K.S.
: TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
: TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
: TITLE: GLUTATHIONE S. TRANSFERASE
: JOURNAL: GENE
: VOLUME:
: ISSUE: 67
: PAGES: 31 - 40
: DATE: 1988
: DOCUMENT NUMBER:
: FILING DATE:
: PUBLICATION DATE:
: RELEVANT RESIDUES IN SEQ ID NO: 13:
: US-08-856-841-13
:
: Query Match: 100.0%; Score 110; DB 3; Length 395;
: Best Local Similarity 100.0%; Pred. No. 1,4e-10;
: Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 1 TYFPNKGTYQYTDQIERPLM 20
: ||||||||||||||||
: Db 252 TYFPNKGTYQYTDQIERPLM 271
:
: RESULT 2
: US-08-856-841-21
: Sequence 21, Application US/08856841
: Patent No. 6274307
: GENERAL INFORMATION:

```

APPLICANT: ERWIN SOUTSCHER
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 398
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 21:
US-08-856-841-21

Query Match 100.0%; Score 110; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPEPKGTQOYTDIERPLM 20
|||||
Db 241 TYPEPKGTQOYTDIERPLM 260

RESULT 3

US-08-856-841-20
Sequence 20, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 415
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIONUM)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: COSSART, Y. E.
AUTHORS: FIELD, A. M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 20:
US-08-856-841-20

Query Match 100.0%; Score 110; DB 3; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPEPKGTQOYTDIERPLM 20
|||||
Db 252 TYPEPKGTQOYTDIERPLM 271

RESULT 4

US-08-856-841-22
Sequence 22, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 543
TYPE: AMINO ACID

TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match 100.0%; Score 110; DB 3; Length 543;
Best Local Similarity 100.0%; Pred. No. 2,1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDIERPLM 20
DB 400 TYFPNKGTOQYTDIERPLM 419

RESULT 5
US-09-438-268-4
Sequence 4, Application US/09438268
Patent No. 6431907
GENERAL INFORMATION:
APPLICANT: Rabinowitz, Joseph E.
APPLICANT: Samulski, Richard J.
APPLICANT: Xiao, Weidong
TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
TITLE OF INVENTION: THE SAME
FILE REFERENCE: 5470-186
CURRENT APPLICATION NUMBER: US/09/438,268
CURRENT FILING DATE: 1999-11-10
EARLIER APPLICATION NUMBER: 60/107,840
EARLIER FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 60/123,651
EARLIER FILING DATE: 1999-03-10
NUMBER OF SEQ ID NOS: 59
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 4
LENGTH: 756
TYPE: PRT
ORGANISM: VITUS
US-09-438-268-4

Query Match 100.0%; Score 110; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TYFPNKGTOQYTDIERPLM 20
|||||

DB 593 TYFPNKGTOQYTDIERPLM 612
RESULT 6
US-09-134-001C-3701
Sequence 3701, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3701
LENGTH: 383
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3701

Query Match 40.9%; Score 45; DB 4; Length 383;
Best Local Similarity 40.0%; Pred. No. 12;
Matches 12; Conservative 1; Mismatches 5; Indels 12; Gaps 1;

QY 2 YFPNKGTOQ-----YTDIERPL 19
DB 212 YGPKGAQKQKMI PKLDSALRHYHDKIEREL 241

RESULT 7
US-09-345-236B-135
Sequence 135, Application US/09345236B
Patent No. 6521454
GENERAL INFORMATION:
APPLICANT: Becnel, James J.
APPLICANT: Tuku, Fukuda
APPLICANT: Moser, Bettina
APPLICANT: Cockburn, Andrew
APPLICANT: White, Susan E.
APPLICANT: Undeen, Albert H.
TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal
TITLE OF INVENTION: Compositions, and Methods for Control of Invertebrates
FILE REFERENCE: 21042.0004
CURRENT APPLICATION NUMBER: US/09/345,236B
CURRENT FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 148
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 135
LENGTH: 119
TYPE: PRT
ORGANISM: mosquito baculovirus
US-09-345-236B-135

Query Match 40.0%; Score 44; DB 4; Length 119;
Best Local Similarity 50.0%; Pred. No. 4.5;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDIE 16
DB 88 TCFPSTGTSEYTSWIK 103

RESULT 8
US-09-345-236B-108
Sequence 108, Application US/09345236B
Patent No. 6521454
GENERAL INFORMATION:
APPLICANT: Becnel, James J.

APPLICANT: Tuku, Fukuda
APPLICANT: Moser, Bettina
APPLICANT: Cockburn, Andrew
APPLICANT: White, Susan E.
APPLICANT: Undeen, Albert H.
TITLE OF INVENTION: No. 6521454e1 Baculoviruses, Insecticidal
TITLE OF INVENTION: Compositions, and Methods for Control of Invertebrates
FILE REFERENCE: 21042.0004
CURRENT APPLICATION NUMBER: US/09/345,236B
CURRENT FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 148
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 108
LENGTH: 235
TYPE: PRT
ORGANISM: mosquito baculovirus
US-09-345-236B-108

Query Match 40.0%; Score 44; DB 4; Length 235;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 TYPKGTQQTDDT 16
DB 204 TCFPTGTSEYTSWIK 219

RESULT 9
US-08-677-862-2
Sequence 2, Application US/08677862
Patent No. 5874230
GENERAL INFORMATION:
APPLICANT: SONG, Ho Yeong
APPLICANT: ROTH, Mike
TITLE OF INVENTION: TRAF2-Associated Kinase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, 34th floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677,862
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: T96-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
TELEFAX: 415 398-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-677-862-2

Query Match 39.5%; Score 43.5; DB 2; Length 763;
Best Local Similarity 47.6%; Pred. No. 50;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
QY 4 PNKGTQ-----YTDQIERPL 19

DB 46 PNISDQVSALSYPDIOQPL 66

RESULT 10
US-09-252-571-2
Sequence 2, Application US/09252571
Patent No. 5981250
GENERAL INFORMATION:
APPLICANT: SONG, Ho Yeong
APPLICANT: ROTH, Mike
TITLE OF INVENTION: TRAF2-Associated Kinase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, 34th floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/252,571
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/677,862
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: T96-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
TELEFAX: 415 398-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-252-571-2

Query Match 39.5%; Score 43.5; DB 2; Length 763;
Best Local Similarity 47.6%; Pred. No. 50;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 4 PNKGTQ-----YTDQIERPL 19
DB 46 PNISDQVSALSYPDIOQPL 66

RESULT 11
US-09-434-065-2
Sequence 2, Application US/09434065
Patent No. 6107074
GENERAL INFORMATION:
APPLICANT: SONG, Ho Yeong
APPLICANT: ROTH, Mike
TITLE OF INVENTION: TRAF2-Associated Kinase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, 34th floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/434,065
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/677,862
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: T96-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
TELEFAX: 415 398-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-434-065-2

Query Match 39.5%; Score 43.5; DB 3; Length 763;
Best Local Similarity 47.6%; Pred. No. 50;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 4 PNKGTOO-----YTDQIERPL 19
DB 46 PNISDQVSALSYSDIQOPL 66

RESULT 12
US-08-789-275-4
Sequence 4, Application US/08789275A
Patent No. 6251664
GENERAL INFORMATION:
APPLICANT: Pallegja, Estivill Xavier
APPLICANT: Pritchard, Melanie
APPLICANT: Vilario, Jordi Guimera
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
Patent No. 6251664
TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
TITLE OF INVENTION: Threonine Protein Kinase (MNB), Expressed in the
FILE REFERENCE: U 011114-4
CURRENT APPLICATION NUMBER: US/08/789,275A
CURRENT FILING DATE: 1997-01-28
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 4
LENGTH: 763
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
US-08-789-275-4

Query Match 39.5%; Score 43.5; DB 3; Length 763;
Best Local Similarity 47.6%; Pred. No. 50;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 4 PNKGTOO-----YTDQIERPL 19
DB 46 PNISDQVSALSYSDIQOPL 66

RESULT 13
US-08-789-275-5
Sequence 5, Application US/08789275A

Patent No. 6251664
GENERAL INFORMATION:
APPLICANT: Pallegja, Estivill Xavier
APPLICANT: Pritchard, Melanie
APPLICANT: Vilario, Jordi Guimera
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
Patent No. 6251664
TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
TITLE OF INVENTION: Threonine Protein Kinase (MNB), Expressed in the
FILE REFERENCE: U 011114-4
CURRENT APPLICATION NUMBER: US/08/789,275A
CURRENT FILING DATE: 1997-01-28
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 5
LENGTH: 763
TYPE: PRT
ORGANISM: Rat norvegicus
FEATURE:
US-08-789-275-5

Query Match 39.5%; Score 43.5; DB 3; Length 763;
Best Local Similarity 47.6%; Pred. No. 50;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 4 PNKGTOO-----YTDQIERPL 19
DB 46 PNISDQVSALSYSDIQOPL 66

RESULT 14
US-07-731-157A-2
Sequence 2, Application US/07731157A
Patent No. 5457032
GENERAL INFORMATION:

APPLICANT: Quax, Wilhelmus J.
APPLICANT: Misset, Onno
APPLICANT: Van der laan, Jan M.
APPLICANT: Lenting, Herman B.M.
TITLE OF INVENTION: Mutated beta-lactam acylase genes
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSER: COOLEY GODWARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/731,157A
FILING DATE: 19910509
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90200962
FILING DATE: 18-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: RAE-VENTER PH.D., BARBARA
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: GBRO-027/00US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-7622
TELEFAX: 415-857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

MOLECULE TYPE: protein
US-07-731-157A-2

Query Match 39.1%; Score 43; DB 1; Length 720;
Best Local Similarity 72.7%; Pred. No. 56;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQCYTDQIER 17
|||:|||||
Db 680 GTTHYSQDIER 690

RESULT 15

US-08-541-780-2
Sequence 2, Application US/08541780
Patent No. 5935831

GENERAL INFORMATION:
APPLICANT: Quax, Wilhelmus J.

APPLICANT: Misset, Onno

APPLICANT: Van der Laan, Jan M.

APPLICANT: Lenting, Herman B.M.

TITLE OF INVENTION: Mutated beta-lactam acylase genes

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:
ADDRESS: COOLEY GODWARD CASTRO HUDDLESON & TATUM

STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/541,780

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/731,157

FILING DATE:

APPLICATION NUMBER: EP 90200962

FILING DATE: 18-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: RAB-VENTER PH.D., BARBARA

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET INFORMATION: GBNO-027/0005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-494-7622

TELEFAX: 415-857-0663

TELEX: 380816 COOLEY PA

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 720 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match 39.1%; Score 43; DB 2; Length 720;
Best Local Similarity 72.7%; Pred. No. 56;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQCYTDQIER 17
|||:|||||

Db 680 GTTHYSQDIER 690

RESULT 16

US-09-252-991A-17112
Sequence 17112, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107156.136
CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 17112

LENGTH: 655

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17112

Query Match 38.2%; Score 42; DB 4; Length 655;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 YFNGKGTQCYTDQIERPL 19
|||:|||||
Db 464 YLPEPGQAQYRDPGEIPL 481

RESULT 17

US-08-447-464-3
Sequence 3, Application US/08447464
Patent No. 5840842

GENERAL INFORMATION:

APPLICANT: Schliesinger, Joseph

APPLICANT: Yan, Hai

TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN

TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESS: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,464

FILING DATE: 24-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/130,570

FILING DATE: 01-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET INFORMATION: 7683-043

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1501 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match

38.2%; Score 42; DB 2; Length 1501;

Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YPNKGTQOY 11
|:|:|:|:|

DB 1057 YPNKGTETY 1066

RESULT 18
US-08-716-679-3

Sequence 3, Application US/08716679
Patent No. 5846800
GENERAL INFORMATION:
APPLICANT: Schlensing, Joseph
APPLICANT: Yan, Hai
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,679
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/130,570
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNTE

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1501 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-716-679-3

Query Match 38.2%; Score 42; DB 2; Length 1501;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YPNKGTQOY 11
|:|:|:|:|

DB 1057 YPNKGTETY 1066

RESULT 19
US-08-348-006B-5

Sequence 5, Application US/08348006B
Patent No. 5658756
GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: J. MARK HAND
STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
CITY: RAHWAY

STATE: NEW JERSEY
COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/348,006B

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/122,032

FILING DATE: 14-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: HAND, J. MARK

REGISTRATION NUMBER: 36,545

REFERENCE/DOCKET NUMBER: 189921A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 908-594-3905

TELEFAX: 908-594-4720

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1911 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-348-006B-5

Query Match 38.2%; Score 42; DB 1; Length 1911;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YPNKGTQOY 11
|:|:|:|:|

DB 1467 YPNKGTETY 1476

RESULT 20
US-08-800-825A-5

Sequence 5, Application US/08800825A
Patent No. 5866397
GENERAL INFORMATION:

APPLICANT: RODAN, GIDEON A.

APPLICANT: SCHMIDT, AZRIEL

APPLICANT: RUTLEDGE, SU JANE

TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN

TITLE OF INVENTION: TYROSINE PHOSPHATASE

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: J. MARK HAND - MERCK & CO., INC.

STREET: 126 E. LINCOLN AVE., P.O. BOX 2000

CITY: RAHWAY

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/800,825A

FILING DATE: 14-SEP-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: HAND, J. MARK

REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 18992DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-3905
TELEFAX: 732-594-4720
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-825A-5

Query Match 38.2%; Score 42; DB 2; Length 1911;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

2 YFPNKGTOOY 11
1467 YWPNRGTETY 1476

RESULT 21
US-09-158-657-5
Sequence 5, Application US/09158657
Patent No. 6214564
GENERAL INFORMATION:
APPLICANT: RODAN, GIDRON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: J. MARK HAND - MERCK & CO., INC.
STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/158.657
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/800,825
FILING DATE: 14-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: HAND, J. MARK
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 18992DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-3905
TELEFAX: 732-594-4720
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-158-657-5

Query Match 38.2%; Score 42; DB 3; Length 1911;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YFPNKGTOOY 11
Db 1467 YWPNRGTETY 1476

RESULT 22
PCT-US94-10166-5
Sequence 5, Application PC/TUS9410166
GENERAL INFORMATION:
APPLICANT: RODAN, GIDRON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: JOHN W. WALLIN III
STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10166
FILING DATE: 09-SEP-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,032
FILING DATE: 14-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: WALLIN, JOHN W III
REGISTRATION NUMBER: 35403
REFERENCE/DOCKET NUMBER: 18992
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
TELEX: 138825
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-10166-5

Query Match 38.2%; Score 42; DB 5; Length 1911;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YFPNKGTOOY 11
Db 1467 YWPNRGTETY 1476

RESULT 23
US-09-095-855-203
Sequence 203, Application US/09095855
Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
Treatment and Diagnosis of Mycobacterial Infections
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:

ADDRESSER: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000,1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 203:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-095-855-203

Query Match 36.4%; Score 40; DB 3; Length 187;
Best Local Similarity 58.3%; Pred. No. 37;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTQYTDQIR 16
DB 161 DKSTHQTNDID 172

LT 24
US-09-205-426-203
Sequence 203, Application US/09205426
Patent No. 6406704
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
TITLE OF INVENTION: Compounds and Methods for Treatment and
FILE REFERENCE: 11000,1002c4
CURRENT APPLICATION NUMBER: US/09/205,426
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: 09/095,855
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 08/997,362
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/873,970
EARLIER FILING DATE: 1997-06-12
EARLIER APPLICATION NUMBER: 08/705,347
EARLIER FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 203
LENGTH: 187
TYPE: PRT

ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: UNSURE
LOCATION: (186)...(186)
US-09-205-426-203

Query Match 36.4%; Score 40; DB 4; Length 187;
Best Local Similarity 58.3%; Pred. No. 37;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTQYTDQIR 16
DB 161 DKSTHQTNDID 172

RESULT 25
US-09-252-991A-22579
Sequence 22579, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22579
LENGTH: 1097
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22579

Query Match 36.4%; Score 40; DB 4; Length 1097;
Best Local Similarity 46.7%; Pred. No. 3e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTQYTDQIR 15
DB 769 TEPFRGQOQOGEQV 783

RESULT 26
US-09-220-641-3
Sequence 3, Application US/09220641
Patent No. 6210923
GENERAL INFORMATION:
APPLICANT: Lee, Cheng-Chi
APPLICANT: Sun, Zhong-Sheng
APPLICANT: Albrecht, Urs
APPLICANT: Bichele, Gregor
TITLE OF INVENTION: Mammalian Circadian Regulator M-RIGU2 (M-PER2)
FILE REFERENCE: D6067
CURRENT APPLICATION NUMBER: US/09/220,641
CURRENT FILING DATE: 1998-12-24
EARLIER APPLICATION NUMBER: US 60/068,886
EARLIER FILING DATE: 1997-12-26
NUMBER OF SEQ ID NOS: 5
SEQ ID NO 3
LENGTH: 1257
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
OTHER INFORMATION: m-rigu2 predicted amino acid sequence.
US-09-220-641-3

Query Match 36.4%; Score 40; DB 3; Length 1257;
Best Local Similarity 47.1%; Pred. No. 3.5e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/06/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oseman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1893 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-188-582-11

Query Match 36.4%; Score 40; DB 1; Length 1893;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 PNKGTOQYTDQIERPL 19
DB 990 PNKPTQKDKKRPV 1005

RESULT 30
US-08-646-715-11
Sequence 11, Application US/08646715
Patent No. 5637686
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 09-MAY-1996
APPLICATION NUMBER: US/08/646,715
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Oseman, Richard A

REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1893 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-646-715-11

Query Match 36.4%; Score 40; DB 1; Length 1893;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 PNKGTOQYTDQIERPL 19
DB 990 PNKPTQKDKKRPV 1005

RESULT 31
US-09-732-210-1674
Sequence 1674, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Miltanck, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-Fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 1674
LENGTH: 151
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-732-210-1674

Query Match 35.9%; Score 39.5; DB 4; Length 151;
Best Local Similarity 52.6%; Pred. No. 34;
Matches 10; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 2 YFPNKGTOYTDQIERPLM 20
DB 59 YFPNKHVHOOI--IKAPLV 74

RESULT 32
US-09-732-210-1672
Sequence 1672, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Miltanck, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-Fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07

```

Sequence 3 Application PC/TUS9506556
GENERAL INFORMATION:
APPLICANT: Cully, Doris F.
APPLICANT: Arena, Joseph P.
APPLICANT: Liu, Ken K.
APPLICANT: Vassiliadis, Demetrios
TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
CHANNELS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wallen, John W.
STREET: 126 E. Lincoln Ave., P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06556
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/249,112
FILING DATE: 25-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wallen, John W.
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 19194
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-06556-3

Query Match 35.9%; Score 39.5; DB 5; Length 510;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 1; Gaps 1.

Cy 1 TYPFNKGTQYTDQIERP 18
|:|:|:|:|:|:|:|:|:|:|
Db 167 TFFPNE-KQAYKHTIDKP 183

RESULT 35
US-08-931-220-2
Sequence 2, Application US/08931220
Patent No. 6030835
GENERAL INFORMATION:
APPLICANT: Musser M.D., James M.
APPLICANT: Kapur M.D., Vivek
TITLE OF INVENTION: Methods and Compositions for Identifying
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEIL, GOTSCHAL & MANGES
STREET: 2882 Sand Hill Road, Suite 280
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025-7022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,220
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,542
FILING DATE: 14-SEP-1994
APPLICATION NUMBER: US 08/160,965
FILING DATE: 02-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter Ph.D., Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: BAYL-004/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6200
TELEFAX: (415) 854-3713
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
POSITION IN GENOME:
MAP POSITION: 203
US-08-931-220-2

Query Match 35.5%; Score 39; DB 3; Length 16;
Best Local Similarity 60.0%; Pred. No. 2.9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 FPNKGTQOQYT 12
:||||:|
Db 4 YPNKGLKDYT 13

RESULT 36
PCT-US95-11723-2
; Sequence 2, Application PC/TUS9511723
; GENERAL INFORMATION:
; APPLICANT: Musser M.D., James M.
; TITLE OF INVENTION: Methods and Compositions for
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAW OFFICES OF BARBARA RAE-VENTER
; STREET: P.O. Box 60039
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11723
; FILING DATE: 14-SEP-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,542
; FILING DATE: 14-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rae-Venter Ph.D., Barbara

REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: BAYL-004/03US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6205
TELEFAX: (415) 424-8760
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
POSITION IN GENOME:
MAP POSITION: 203
PCT-US95-11723-2

Query Match 35.5%; Score 39; DB 5; Length 16;
Best Local Similarity 60.0%; Pred. No. 2.9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 FPNKGTQOQYT 12
:||||:|
Db 4 YPNKGLKDYT 13

RESULT 37
PCT-US96-05997-15
; Sequence 15, Application PC/TUS9605997
; GENERAL INFORMATION:
; APPLICANT: Musser M.D., James M.
; APPLICANT: Kapur M.D., Vivek
; APPLICANT: Ananthaaswamy, H. N.
; APPLICANT: Fernandez, A.
; TITLE OF INVENTION: Use of extracellular cysteine protease
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WEIL, GOTSCHAL & MANGES
; STREET: 2882 Sand Hill Road, Suite 280
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025-7022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05997
; FILING DATE: 01-MAY-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,542
; FILING DATE: 14-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/160,965
; FILING DATE: 02-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rae-Venter Ph.D., Barbara
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: BAYL-004/02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 926-6200
; TELEFAX: (415) 854-3713
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
POSITION IN GENOME:
MAP POSITION: 203
PCT-US96-05997-15

Query Match 35.5%; Score 39; DB 5; Length 16;
Best Local Similarity 60.0%; Pred. No. 2.9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PNRKGTQOYTQ 12
4 YPNKGLKDYT 13

RESULT 38
US-09-107-532A-4773

Sequence 4773, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESS: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Atinello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 4773:

SEQUENCE CHARACTERISTICS:

LENGTH: 124 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...124

SEQUENCE DESCRIPTION: SEQ ID NO: 4773:

US-09-107-532A-4773

Query Match 35.5%; Score 39; DB 4; Length 124;
Best Local Similarity 58.3%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 PNRKGTQOYTQ 14
64 PNRGVGVGYVSO 75

RESULT 39
US-08-311-731A-299

Sequence 299, Application US/08311731A
Patent No. 6583266

GENERAL INFORMATION:

APPLICANT: SMITH, DOUGLAS

APPLICANT: MAO, JEN-I

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR

NUMBER OF SEQUENCES: 411

CORRESPONDENCE ADDRESS:

ADDRESS: WOLF, GREENFIELD & SACKS, P.C.

STREET: 600 ATLANTIC AVENUE

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: USA

ZIP: 02210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/311,731A

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: GATES, EDWARD R.

REGISTRATION NUMBER: 31,616

REFERENCE/DOCKET NUMBER: C0044/7125

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/720-2441

TELEFAX: 617/720-3500

INFORMATION FOR SEQ ID NO: 299:

SEQUENCE CHARACTERISTICS:

LENGTH: 166 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Mycobacterium leprae

US-08-311-731A-299

Query Match 35.5%; Score 39; DB 4; Length 166;
Best Local Similarity 41.2%; Pred. No. 47;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PNRKGTQOYTQIERPLM 20
35 PNRGTDTDFDNRPLV 51

RESULT 40
US-09-134-001C-5493

Sequence 5493, Application US/09134001C
Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO: 5493
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5493

Query Match 35.5%; Score 39; DB 4; Length 241;
Best Local Similarity 63.6%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YFPNKGTOQYT 12
|||:|
DB 223 YFNGKGKORYT 233

Search completed: August 20, 2003, 09:34:31
Job time : 22.0619 secs

; Publication No. US20030143248A1

```
GENERAL INFORMATION:
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: BARNETT, Susan W.
APPLICANT: LIAN, Ying
APPLICANT: ENGELBRECHT, Susan
APPLICANT: VAN RENSBURG, Estrelita J.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
FILE REFERENCE: P18133.003 / 2302-18133
CURRENT APPLICATION NUMBER: US/10/190,435
CURRENT FILING DATE: 2002-12-30
NUMBER OF SEQ ID NOS: 319
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 869
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: TV1_8_5
10-190-435-4
Query Match          40.0%; Score 44; DB 12; Length 869;
Best Local Similarity 36.8%; Pred. No. 1e+02;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 TYPNKGTOOYTDQIERPL 19
DB 399 TYPNKGTYKYNNGNSSLPI 417

RESULT 3
US-09-864-761-34120
Sequence 34120, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmlca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonex Sequence Listing Engine vers. 1.1
SEQ ID NO 34120
LENGTH: 50
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000154.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
OTHER INFORMATION: SWISSPROT HIT: Q13627, EVALUATE 9.00e-24
OTHER INFORMATION: EST_HUMAN HIT: BE817813.1, EVALUATE 1.00e-22
US-09-864-761-34120

Query Match          39.5%; Score 43.5; DB 9; Length 50;
Best Local Similarity 47.6%; Pred. No. 4.8;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 4 PNKGTQO-----YTDQIERPL 19
DB 27 PNISDQVSALYSDDIQDPL 47

RESULT 4
US-09-738-626-3714
Sequence 3714, Application US/09738626
Patent No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 3714
LENGTH: 661
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-3714

Query Match          39.1%; Score 43; DB 10; Length 661;
Best Local Similarity 36.8%; Pred. No. 1.1e+02;
```


Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0

OY 1 TYFPNKGTYQYTDDIERPL 19
::|::||:|:
Db 542 SYTPRGYGQAFTSHANPL 560

```

RESULT 5
US-09-801-574-65
: Sequence 65, Application US/09801574
: Patent No. US20020081592A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Peijing Jeremy
: APPLICANT: Page, David C.
: TITLE OF INVENTION: Reproduction-Specific Genes
: FILE REFERENCE: 0399 2007-002
: CURRENT APPLICATION NUMBER: US/09/801,574
: CURRENT FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: 60/187,518
: PRIOR FILING DATE: 2000-03-07
: PRIOR APPLICATION NUMBER: 60/261,557
: PRIOR FILING DATE: 2001-01-12
: NUMBER OF SEQ ID NOS: 90
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 65
: LENGTH: 913
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-801-574-65

```

Query Match	39.1%;	Score 43;	DB 9;	Length 913;
Best Local Similarity	50.0%;	Pred. No. 1.5e+02;		
Matches	8;	Conservative	2;	Mismatches 6;
			Indels	0;
			Gaps	0;

```

RESULT 6
US-10-156-761-10200
: Sequence 10200, Application US/10156761
: Publication No. US20030119018A1
GENERAL INFORMATION:
: APPLICANT: OMURA, SATOSHI
: APPLICANT: IKEDA, HARUO
: APPLICANT: ISHIKAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIBA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHITUKU
: APPLICANT: HATTORI, MASAHIRA
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-262
: CURRENT FILING DATE: 2002-05-29
: PRIOR APPLICATION NUMBER: US/10/156, 761
: PRIOR FILING DATE: 2001-05-30
: PRIOR FILING DATE: 2001-05-30
: PRIOR APPLICATION NUMBER: JP 2001-272697
: PRIOR FILING DATE: 2001-08-02
: NUMBER OF SEQ ID NOS: 15109
: SEQ ID NO 10200
: LENGTH: 437
: TYPE: PRT
: ORGANISM: Streptomyces avermitilis
: US-10-156-761-10200

```

Query Match	38.2%;	Score 42;	DB 15;	Length 437;
Best Local Similarity	66.7%;	Pred. No. 97;		
Matches	8;	Conservative	0;	Mismatches 4;
				Indels 0;
				Gaps 0;

```

RESULT: 7
US-09-808-602-54
; Sequence 54, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Heriman, John L
; APPLICANT: Majumder, Kuntud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 1502
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-54

```

Query Match	38.2%	Score 42;	DB 10;	Length 1502;
Best Local Similarity	60.0%;	Pred. No. 3.9e+02;		
Matches	6;	Conservative	3;	Mismatches 1; Indels 0; Gaps 0;

```

RESULT 8
US-09-800-198-44
Sequence 44, Application US/09800198
Publication No. US20030087816A1
GENERAL INFORMATION:
APPLICANT: Verneet, Cornie AM
APPLICANT: Fernandes, Elma
APPLICANT: Shinkets, Richard A
APPLICANT: Hermann, John L
APPLICANT: Majumder, Kumud
APPLICANT: Mishra, Vishna
APPLICANT: Mezes, Peter S
APPLICANT: Rastelli, Luca
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-697
CURRENT APPLICATION NUMBER: US/09/800,198
CURRENT FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 44
LENGTH: 1502
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-198-44

```

Query Match	38.2%	Score 42;	DB 11;	Length 1502;
Best Local Similarity	60.0%	Pred. No. 3.9e+02;		
Matches	6;	Conservative	3;	Mismatches 1;
				Indels 0;
				Gaps 0;
QY	2 YFPNKCTQY 11			
	: : : :			

Db 1058 YWPNRGTEY 1067

RESULT 9

US-09-808-602-55
; Sequence 55, Application US/09808602
; Patent No. US2002015515A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandez, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: NO. US002015515A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 55
; LENGTH: 1948
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-55

Query Match 38.2%; Score 42; DB 10; Length 1948;
Best Local Similarity 60.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YFPNKGTOOY 11
|:|:|:|:
Db 1504 YWPNRGTEY 1513

RESULT 10

US-09-800-198-45
; Sequence 45, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1948
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-198-45

Query Match 38.2%; Score 42; DB 11; Length 1948;
Best Local Similarity 60.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YFPNKGTOOY 11

Db 1504 YWPNRGTEY 1513

RESULT 11

US-09-815-242-5346
; Sequence 5346, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5346
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5346

Query Match 37.3%; Score 41; DB 9; Length 266;
Best Local Similarity 52.4%; Pred. No. 81;
Matches 11; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

Qy 2 YFPNKGTO--QYTDQIERPLM 20
|:|:|:|:|:|:|:
Db 41 YGPRGEQVKQYTVIARKLM 61

RESULT 12

US-09-815-242-12340
; Sequence 12340, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078

PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12340
LENGTH: 266
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12340

Query Match 37.3%; Score 41; DB 9; Length 266;
Best Local Similarity 52.4%; Pred. No. 81;
Matches 11; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

Oy 2 YFPNKGTO-QYTDQIERPLM 20
Db 41 YGPRGEQVKQYTVIARKLM 61

RESULT 13
US-09-815-242-12989
Sequence 12989, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12989
LENGTH: 266
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12989

Query Match 37.3%; Score 41; DB 9; Length 266;
Best Local Similarity 52.4%; Pred. No. 81;
Matches 11; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

Oy 2 YFPNKGTO-QYTDQIERPLM 20
Db 41 YGPRGEQVKQYTVIARKLM 61

RESULT 14
US-09-815-242-13476
Sequence 13476, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13476
LENGTH: 283
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13476

Query Match 37.3%; Score 41; DB 9; Length 283;
Best Local Similarity 46.7%; Pred. No. 87;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 1 TYFPNKGTOQYTDQI 15
Db 205 TSVPKNSQYTYETV 219

RESULT 15
US-10-238-075-691
Sequence 691, Application US/10238075
Publication No. US20030148324A1
GENERAL INFORMATION:
APPLICANT: I.N.S.E.R.M.
TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolat-
FILE REFERENCE: BLANDINE
CURRENT APPLICATION NUMBER: US/10/238,075
CURRENT FILING DATE: 2002-09-10
PRIOR APPLICATION NUMBER: 0003145
PRIOR FILING DATE: 2000-03-10
NUMBER OF SEQ ID NOS: 1576
SOFTWARE: Patentin version 3.1
SEQ ID NO 691
LENGTH: 605
TYPE: PRT
ORGANISM: Escherichia coli

US-10-238-075-691

Query Match 37.3%; Score 41; DB 12; Length 605;
Best Local Similarity 44.4%; Pred. No. 2.1e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 YFPNKGTOQYTDQIERPL 19
Db 282 YFPNENSPFYTLKIDSAL 299

RESULT 16

US-10-205-979-49
; Sequence 49, Application US/10205979
; Publication No. US20030147861A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; APPLICANT: Abernethy, Nevlin
; TITLE OF INVENTION: Compounds and Methods for the Modulation
; FILE REFERENCE: 11000.1063U
; CURRENT APPLICATION NUMBER: US/10/205,979
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/308,446
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(187)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-205-979-49

Query Match 36.4%; Score 40; DB 12; Length 187;
Best Local Similarity 58.3%; Pred. No. 80;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 NKGTOQYTDQIE 16
Db 161 DKSTHQYTNQID 172

RESULT 17

US-051-643-203
; Sequence 203, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; FILE REFERENCE: 11000.1008C2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 203
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
; NAME/KEY: UNSURE
; LOCATION: (186)...(186)

US-10-051-643-203

Query Match 36.4%; Score 40; DB 14; Length 187;
Best Local Similarity 58.3%; Pred. No. 80;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 NKGTOQYTDQIE 16
Db 161 DKSTHQYTNQID 172

RESULT 18

US-10-029-180-78
; Sequence 78, Application US/10029180
; Publication No. US20020182708A1
; GENERAL INFORMATION:
; APPLICANT: Call, Brian M.
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin T.
; APPLICANT: Milna, G. Todd
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeffrey C.
; APPLICANT: Trueheart, Josh
; APPLICANT: Zhang, Lixin
; TITLE OF INVENTION: No. US20020182708A1 Regulators of Fungal Gene Expression
; FILE REFERENCE: MIC-004
; CURRENT APPLICATION NUMBER: US/10/029,180
; CURRENT FILING DATE: 2001-12-22
; PRIOR APPLICATION NUMBER: US 60/257,431
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fungal gene
US-10-029-180-78

Query Match 36.4%; Score 40; DB 14; Length 405;
Best Local Similarity 44.4%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 2 YFPNKGTOQYTDQIERPL 19
Db 180 YFANKADSPVDPSPILPL 197

RESULT 19

US-10-029-180-112
; Sequence 112, Application US/10029180
; Publication No. US20020182708A1
; GENERAL INFORMATION:
; APPLICANT: Call, Brian M.
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin T.
; APPLICANT: Milna, G. Todd
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeffrey C.
; APPLICANT: Trueheart, Josh
; APPLICANT: Zhang, Lixin
; TITLE OF INVENTION: No. US20020182708A1 Regulators of Fungal Gene Expression
; FILE REFERENCE: MIC-004
; CURRENT APPLICATION NUMBER: US/10/029,180
; CURRENT FILING DATE: 2001-12-22
; PRIOR APPLICATION NUMBER: US 60/257,431
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112
; LENGTH: 434
; TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: fungal gene
US-10-029-180-112

Query Match 36.4%; Score 40; DB 14; Length 434;
Best Local Similarity 44.4%; Pred. No. 2.1e+02;
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 YFPNGTQGYTDQIERPL 19
|||:|:
Db 180 YFANGADSFVDPISLPL 197

RESULT 20
US-09-815-242-13218
Sequence 13218, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
PROKARYOTES
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
ID NO 13218
LENGTH: 436
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13218

Query Match 36.4%; Score 40; DB 9; Length 436;
Best Local Similarity 43.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 PNKGTQGYTDQIERPL 19
|||:|:
Db 355 POKGTBYMSKILRL 370

RESULT 21
US-10-097-065-139
Sequence 139, Application US/10097065
Publication No. US20030055236A1
GENERAL INFORMATION:
APPLICANT: Moore, Paul A. et al.
TITLE OF INVENTION: 110 Human Secreted Proteins
FILE REFERENCE: P2021P1
CURRENT APPLICATION NUMBER: US/10/097,065
CURRENT FILING DATE: 2002-03-14

PRIOR APPLICATION NUMBER: PCT/US98/27059
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: 60/070,923
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,007
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,057
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,006
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,369
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,367
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,053
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,064
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,054
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,008
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,365
PRIOR FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 672
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 139
LENGTH: 549
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (132)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (398)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-097-065-139

Query Match 36.4%; Score 40; DB 15; Length 549;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTQGYTDQIERPLM 20
|||:|:
Db 406 NKATQLFNTQTDIPNM 421

RESULT 22
US-10-199-672-326
Sequence 326, Application US/10199672
Publication No. US20030148442A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Qian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/199,672
CURRENT FILING DATE: 2002-07-18

```

; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-199-672-326
```

```

Query Match      36.4%; Score 40; DB 12; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy      5 NKGQYTDQIERPLM 20
Db      414 NKATQFTNQTDIPNM 429
```

```

RESULT 23
US-10-187-749-326
; Sequence 326, Application US/10187749
; Publication No. US20030153036A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/187,749
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-194-457-326
```

```

; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-749-326
```

```

Query Match      36.4%; Score 40; DB 12; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy      5 NKGQYTDQIERPLM 20
Db      414 NKATQFTNQTDIPNM 429
```

```

RESULT 24
US-10-194-457-326
; Sequence 326, Application US/10194457
; Publication No. US20030153037A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C296
; CURRENT APPLICATION NUMBER: US/10/194,457
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-194-457-326
```

```

Query Match      36.4%; Score 40; DB 12; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
```

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 5 NKGQYTDQIERPLM 20
|||:|:|
Db 414 NKATQFTNQDIPNM 429

RESULT 25
US-10-052-586-326

; Sequence 326, Application US/10052586
; Publication No. US20020127584A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Aubelin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3430R1C1

; CURRENT APPLICATION NUMBER: US/10/052,586

; CURRENT FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059266

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/063120

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063121

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063486

; PRIOR FILING DATE: 1997-10-21

; PRIOR APPLICATION NUMBER: 60/063540

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063541

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063544

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063564

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063734

; PRIOR FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/063870

; PRIOR FILING DATE: 1997-10-31

; PRIOR APPLICATION NUMBER: 60/064103

; PRIOR FILING DATE: 1997-10-31

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066120

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/066466

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/066772

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/069335

; PRIOR FILING DATE: 1997-12-11

; PRIOR APPLICATION NUMBER: 60/069425

; PRIOR FILING DATE: 1997-12-12

; PRIOR APPLICATION NUMBER: 60/069870

; PRIOR FILING DATE: 1997-12-17

; PRIOR APPLICATION NUMBER: 60/068017

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: 60/077632

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077649

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/078886

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/078939

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/079664

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/079786

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/080107

; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: 60/080194

; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: 60/080327

; PRIOR FILING DATE: 1998-04-01

; PRIOR APPLICATION NUMBER: 60/080333

; PRIOR FILING DATE: 1998-04-01

; PRIOR APPLICATION NUMBER: 60/081049

; PRIOR FILING DATE: 1998-04-08

; PRIOR APPLICATION NUMBER: 60/081070

; PRIOR FILING DATE: 1998-04-08

; PRIOR APPLICATION NUMBER: 60/081195

; PRIOR FILING DATE: 1998-04-09

; PRIOR APPLICATION NUMBER: 60/081838

; PRIOR FILING DATE: 1998-04-15

; PRIOR APPLICATION NUMBER: 60/082568

; PRIOR FILING DATE: 1998-04-21

; PRIOR APPLICATION NUMBER: 60/082569

; PRIOR FILING DATE: 1998-04-21

; PRIOR APPLICATION NUMBER: 60/082704

; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/082797

; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/083495

; PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/083496

; PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/083499

; PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/083559

; PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/084366

; PRIOR FILING DATE: 1998-05-05

; PRIOR APPLICATION NUMBER: 60/084414

; PRIOR FILING DATE: 1998-05-06

; PRIOR APPLICATION NUMBER: 60/084639

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/084640

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/084643

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/085573

; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/085579

; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/085580

; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/085582

; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/085700

; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/086023

; PRIOR FILING DATE: 1998-05-18

; PRIOR APPLICATION NUMBER: 60/086392

; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: 60/086486

; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: 60/087098

; PRIOR FILING DATE: 1998-05-28

```

; PRIOR APPLICATION NUMBER: 60/087208
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088722
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908

```

```

Query Match      36.4%; Score 40; DB 14; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy      5 NKGTQYTDIETPLM 20
Db      414 NKATQLFTNQTIDIPNM 429

```

```

RESULT 26
US-10-174-590-326
; Sequence 326, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-326

```

```

Query Match      36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy      5 NKGTQYTDIETPLM 20
Db      414 NKATQLFTNQTIDIPNM 429

```

```

RESULT 27
US-10-176-758-326
; Sequence 326, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-326

```

```

Query Match      36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy      5 NKGTQYTDIETPLM 20

```


Db 414 NKATQFTNQTDIPNM 429

RESULT 28

US-10-175-737-326

; Sequence 326, Application US/10175737
; Publication No. US20030013153A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C50

; CURRENT APPLICATION NUMBER: US/10/175,737

; CURRENT FILING DATE: 2002-06-19

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 326

; LENGTH: 557

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-175-737-326

Query Match

Best Local Similarity 36.4%; Score 40; DB 15; Length 557;

Best Local Similarity 50.0%; Pred. No. 2.7e+02;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTQYTDIERPLM 20

Db 414 NKATQFTNQTDIPNM 429

RESULT 29

US-10-173-706-326

; Sequence 326, Application US/10173706
; Publication No. US20030022293A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C7

; CURRENT APPLICATION NUMBER: US/10/173,706

; CURRENT FILING DATE: 2002-06-17

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 326

; LENGTH: 557

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-173-706-326

Query Match

Best Local Similarity 36.4%; Score 40; DB 15; Length 557;

Best Local Similarity 50.0%; Pred. No. 2.7e+02;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTQYTDIERPLM 20

Db 414 NKATQFTNQTDIPNM 429

RESULT 30

US-10-175-738-326

; Sequence 326, Application US/10175738
; Publication No. US20030022294A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C45

; CURRENT APPLICATION NUMBER: US/10/175,738

; CURRENT FILING DATE: 2002-06-19

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 326

; LENGTH: 557

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-175-738-326

Query Match

Best Local Similarity 36.4%; Score 40; DB 15; Length 557;

Best Local Similarity 50.0%; Pred. No. 2.7e+02;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTQYTDIERPLM 20

Db 414 NKATQFTNQTDIPNM 429

RESULT 31

US-10-175-752-326

; Sequence 326, Application US/10175752
; Publication No. US20030022295A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C60

; CURRENT APPLICATION NUMBER: US/10/175,752

; CURRENT FILING DATE: 2002-06-19

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 326

; LENGTH: 557

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-175-752-326

Query Match

Best Local Similarity 36.4%; Score 40; DB 15; Length 557;

Best Local Similarity 50.0%; Pred. No. 2.7e+02;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Query Match 36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTQYTDQIERPLM 20
|||:|:|:
DB 414 NKATQFTNQTDIPNM 429

RESULT 32
US-10-176-482-326

; Sequence 326, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

; ORGANISM: Homo Sapien
US-10-176-757-326

Query Match 36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTQYTDQIERPLM 20
|||:|:|:
DB 414 NKATQFTNQTDIPNM 429

RESULT 34
US-10-176-913-326

; Sequence 326, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

; ORGANISM: Homo Sapien
US-10-176-913-326

Query Match 36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTQYTDQIERPLM 20
|||:|:|:
DB 414 NKATQFTNQTDIPNM 429

RESULT 35
US-10-180-552-326

; Sequence 326, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

FILE REFERENCE: P3430R1C6
TITLE OF INVENTION: ACIDS ENCODING THE SAME
SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

CURRENT APPLICATION NUMBER: US/10/176,757
FILE REFERENCE: P3430R1C6
CURRENT FILING DATE: 2002-06-20

Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 612

SEQ ID NO 326
LENGTH: 557
TYPE: PRT

SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-180-552-326

Query Match 36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTOQYTDQIERPLM 20
|||:|:|:
Db 414 NKATQLFTNQTDIPNM 429

RESULT 36
US-10-180-557-326
Sequence 326, Application US/10180557
Publication No. US20030022301A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C147
CURRENT APPLICATION NUMBER: US/10/180,557

Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326

LENGTH: 557

TYPE: PRT
ORGANISM: Homo Sapien
US-10-180-557-326

Query Match 36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTOYTDQIERPLM 20
|||:|:|:
Db 414 NKATQLFTNQTDIPNM 429

RESULT 37
US-10-173-700-326
Sequence 326, Application US/10173700
Publication No. US20030027262A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C14
CURRENT APPLICATION NUMBER: US/10/173,700

CURRENT FILING DATE: 2002-06-17
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-173-700-326

Query Match 36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTOYTDQIERPLM 20
|||:|:|:
Db 414 NKATQLFTNQTDIPNM 429

RESULT 38
US-10-174-572-326
Sequence 326, Application US/10174572
Publication No. US20030027263A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C40
CURRENT APPLICATION NUMBER: US/10/174,572

Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326

LENGTH: 557

TYPE: PRT
ORGANISM: Homo Sapien
US-10-174-572-326

Query Match 36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTOYTDQIERPLM 20
|||:|:|:
Db 414 NKATQLFTNQTDIPNM 429

RESULT 39
US-10-174-579-326
Sequence 326, Application US/10174579
Publication No. US20030027264A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C31
; CURRENT APPLICATION NUMBER: US/10/174,579
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-579-326

```

```

Query Match          36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches      8; Conservative      3; Mismatches      5; Indels      0; Gaps      0;

```

```

QY      5 NKGTQOYTDQIERPLM 20
DB      414 NKATQLEFTNQTDIPNM 429

```

```

RESULT 40
US-10-174-582-326
; Sequence 326, Application US/10174582
; Publication No. US20030027265A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C36
; CURRENT APPLICATION NUMBER: US/10/174,582
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
0-174-582-326

```

```

Query Match          36.4%; Score 40; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches      8; Conservative      3; Mismatches      5; Indels      0; Gaps      0;

```

```

QY      5 NKGTQOYTDQIERPLM 20
DB      414 NKATQLEFTNQTDIPNM 429

```

Search completed: August 20, 2003, 10:08:19
Job time : 82.2371 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 22.4742 Seconds
(without alignments)
85.581 Million cell updates/sec

Title: US-09-991-433-48

Perfect score: 110

Sequence: 1 TYPNKGTOQYTDQIERPLM 20

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	781	1	VCPV19
2	58	52.7	458	2	B90625
3	55	50.0	459	2	T11333
4	54	49.1	458	2	B90623
5	53	48.2	458	2	B90615
6	53	48.2	459	2	T11202
7	52	47.3	455	2	B90619
8	52	47.3	459	2	T11084
9	51	46.4	458	2	T11528
10	51	46.4	458	2	T11176
11	51	46.4	458	2	B90613
12	50	45.5	458	2	B90627
13	49	44.5	458	2	B90621
14	49	44.5	458	2	T16580
15	48	43.6	1336	2	S41794
16	46	41.8	458	2	B90617
17	46	41.8	459	2	T11031
18	46	41.8	714	2	S76082
19	45.5	41.4	519	2	T45764
20	45	40.9	299	2	A63278
21	45	40.9	868	2	T06827
22	44	40.0	71	2	D87036
23	44	40.0	508	1	VGWNR
24	44	40.0	547	2	JC4519
25	44	40.0	548	2	C64076
26	44	40.0	1173	2	B97208
27	43.5	39.5	754	2	JC4898
28	43	39.1	89	2	I68530
29	43	39.1	177	2	AH0069

30	43	39.1	459	2	T11189	NADH2 dehydrogenas
31	43	39.1	1577	2	T15851	hypothetical prote
32	42	38.2	88	2	A44929	protein-tyrosine-p
33	42	38.2	333	2	D88970	protein F15E11.11
34	42	38.2	398	2	T16540	protein-tyrosine-p
35	42	38.2	514	2	B86164	hypothetical prote
36	42	38.2	538	2	A85075	probable transposo
37	42	38.2	669	2	S14535	asparagine-rich pr
38	42	38.2	765	2	AD0778	beta-glucosidase (
39	42	38.2	859	2	T06429	lipoygenase (EC 1
40	42	38.2	1262	1	B48758	protein-tyrosine-p
41	42	38.2	1496	1	A48758	protein-tyrosine-p
42	42	38.2	1501	2	I58148	protein-tyrosine-p
43	42	38.2	1863	2	S46217	protein-tyrosine-p
44	42	38.2	1907	2	S50893	protein-tyrosine-p
45	41	37.3	206	2	A81231	50S ribosomal prot

ALIGNMENTS

RESULT 1

VCPV19

coat protein VP1 - parvovirus B19 (strain Au)

C:Species: parvovirus B19

A>Note: host Homo sapiens (nan)

C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C:Accession: A24299

R:Shade, R.O., Blundell, M.C., Cotmore, S.F., Tattersall, P., Atwell, C.R.

J. Virol. 58, 921-936, 1986

A>Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr

A:Reference number: A24299; MUID:86200451; PMID:3701931

A:Accession: A24299

A:Molecule type: DNA

A:Residues: 1-781 <SHA>

A:Cross-references: EMBL:M13178; NID:G333375; PID:AAA6867.1; PID:G333377

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein

Query Match

Best Local Similarity 100.0%; Score 110; DB 1; Length 781;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPNKGTOQYTDQIERPLM 20

DB 618 TYPNKGTOQYTDQIERPLM 637

RESULT 2

B90625 NADH dehydrogenase chain 4 ND4 [imported] - Casuarinus casuarinus mitochondrion

C:Species: mitochondrion Casuarinus casuarinus

C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001

C:Accession: B90625

R:Haddad, O.; Baker, A.J.

Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001

A>Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene

A:Reference number: A99613; MUID:21263106; PMID:11370967

A:Accession: B90625

A:Structure: Preliminary

A:Molecule type: DNA

A:Residues: 1-458 <KIR>

A:Cross-references: GB:NC_002778; NID:G14141873; PID:NF_115359.1; GSPDB:GN00162

C:Genetics:

A:Gene: ND4

A:Genome: mitochondrion

A:Genetic code: SGC1

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C:Keywords: mitochondrion

Query Match

Best Local Similarity 52.7%; Score 58; DB 2; Length 458;

Matches 12; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C/Accession: B90627
R/Haddock, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A/Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene
A/Reference number: A99613; MUID:21263106; PMID:11370967
A/Accession: B90627
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-458 <KUR>
A/Cross-references: GB:NC_002772; NID:g14141810; PIDN:NP_115282.1; GSPDB:GN00163
A/Genes: ND4
A/Genetic code: SSCI
A/Suprafamily: NADH dehydrogenase (ubiquinone) chain 4
C/Keywords: mitochondrion

Query Match 45.5%; Score 50; DB 2; Length 458;
Best Local Similarity 50.0%; Pred. No. 3.9;
Matches 11; Conservative 4; Mismatches 5; Indels 2; Gaps 1;
Qy 1 TYFPNKGTOQYT--DQIERPLM 20
: |||: |||: |||:
Db 45 SYFPHKCTQWTSDISAPL 66

RESULT 13
B90621
NADH dehydrogenase chain 4 ND4 [imported] - Tinamus major mitochondrion
C/Species: mitochondrion tinamus major
C/Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C/Accession: B90621
R/Haddock, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A/Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene
A/Reference number: A99613; MUID:21263106; PMID:11370967
A/Accession: B90621
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-458 <KUR>
A/Cross-references: GB:NC_002781; NID:g14141957; PIDN:NP_115398.1; GSPDB:GN00160
C/Genetics:
A/Genes: ND4
A/Genome: mitochondrion
A/Suprafamily: NADH dehydrogenase (ubiquinone) chain 4
C/Keywords: mitochondrion

Query Match 44.5%; Score 49; DB 2; Length 458;
Best Local Similarity 45.5%; Pred. No. 5.6;
Matches 10; Conservative 6; Mismatches 4; Indels 2; Gaps 1;
Qy 1 TYFPNKGTOQYT--DQIERPLM 20
: |||: |||: |||:
Db 45 SYFPHKSSQWSDISAPL 66

RESULT 14
T16580
hypothetical protein K07B12.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T16580
R/Fulton, L.
Submitted to the EMBL Data Library, May 1994
A/Description: The sequence of C. elegans cosmid K07B12.
A/Reference number: Z18540
A/Accession: T16580
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-13055 <FTL>
A/Cross-references: EMBL:U00054; NID:g485140; PID:g485141; PIDN:AAA50715.1; CESP:K07B12.

A/Experimental source: strain Bristol N2
C/Genetics:
A/Genes: CESP:K07B12.1
A/Introns: 46/3; 85/2; 201/1; 278/2; 470/2; 817/1; 927/1; 960/3; 1265/1; 1322/1; 1478/3;
; 6014/3; 6159/3; 6665/2; 7266/3; 7895/3; 8669/3; 8726/3; 9803/3; 10537/3; 12234/2; 1229/
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 4 PKNKTOQYTDQIERP 18
: |||: |||: |||:
Db 1879 PNEWTQYTDQIERP 1893

RESULT 15
S41794
SEC3 protein - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YER008c; PS1 protein
C/Species: Saccharomyces cerevisiae
C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 21-Jul-2000
C/Accession: S41794; S50466; S72238
R/Haarer, B.K.; Petzold, A.S.; Brown, S.S.
Submitted to the EMBL Data Library, July 1993
A/Description: Identification of mutations that are synthetically lethal with altered ye
A/Reference number: S41793
A/Accession: S41794
A/Molecule type: DNA
A/Residues: 1-1336 <HAA>
A/Cross-references: EMBL:L22204; NID:g347714; PID:g347716
R/Dierich, F.S.
submitted to the EMBL Data Library, December 1994
A/Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9455, 9867, and lambda

A/Reference number: S50459
A/Accession: S50466
A/Molecule type: DNA
A/Residues: 1-1336 <DIE>
A/Cross-references: EMBL:U18778; NID:g603592; PID:g603600; MIPS:YER008c
R/Haarer, B.K.; Corbett, A.; Kwon, Y.; Petzold, A.S.; Silver, P.; Brown, S.S.
Genetics 144, 495-510, 1996
A/Title: SEC3 mutations are synthetically lethal with profilin mutations and cause defect
A/Reference number: S72237; MUID:97044444; PMID:8889515
A/Accession: S72238
A/Molecule type: DNA
A/Residues: 1-1336 <HAW>
A/Cross-references: EMBL:L22204; NID:g347714; PIDN:AAA9380.1; PID:g347716
C/Genetics:
A/Genes: SGD:SEC3; PS1
A/Cross-references: MIPS:YER008c; SGD:S0000810
A/Map position: 5R

Query Match 43.6%; Score 48; DB 2; Length 1336;
Best Local Similarity 42.1%; Pred. No. 27;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
Qy 1 TYFPNKGTOYTDQIERP 19
: |||: |||: |||:
Db 498 TYFERSTARAHQDKKFL 516

RESULT 16
B90617
NADH dehydrogenase chain 4 ND4 [imported] - Pterocnemis pennata mitochondrion
C/Species: mitochondrion Pterocnemis pennata
C/Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C/Accession: B90617
R/Haddock, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A/Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene
A/Reference number: A99613; MUID:21263106; PMID:11370967
A/Accession: B90617
A/Status: preliminary
A/Molecule type: DNA

A:Residues: 1-458 <KUR>
 A:Cross-references: GB:NC_002783; NID:G14141901; PIDN:NP_115424.1; GSPDB:GN00158
 C:Genetics:
 A:Gene: ND4
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
 C:Keywords: mitochondrion

Query Match 41.8%; Score 46; DB 2; Length 458;
 Best Local Similarity 45.5%; Pred. No. 17;
 Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 1 YFPNKGTOOYT--DOIERPIM 20
 |||:|:|:|:|:
 DB 45 YFPNKGTONSVWAGIDQISAPLL 66

RESULT 17

031

A:Species: mitochondrion *Aythya americana* (redhead)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
 C:Accession: T11031
 R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
 Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
 A:Title: Multiple independent origins of mitochondrial gene order in birds.
 A:Reference number: Z17242
 A:Accession: T11031
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-459 <JOH>
 A:Cross-references: EMBL:AF090337; NID:G4887659; PID:G4887669; PIDN:AAD22261.1
 C:Genetics:
 A:Genome: mitochondrion
 A:Mobile element: SGC1
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
 C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 41.8%; Score 46; DB 2; Length 459;
 Best Local Similarity 45.5%; Pred. No. 17;
 Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 1 YFPNKGTOOYT--DOIERPIM 20
 |||:|:|:|:|:
 DB 45 YFPNKGTONSVWAGIDQISAPLL 66

Query Match 41.8%; Score 46; DB 2; Length 459; Best Local Similarity 45.5%; Pred. No. 17; Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 1 YFPNKGTOOYT--DOIERPIM 20
 |||:~|:~|:~|:~|:
 DB 45 YFPNKGTONSVWAGIDQISAPLL 66

A:Species: *Synechocystis* sp. (strain PCC 6803)
 C:Species: *Synechocystis* sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 02-Aug-2002
 C:Accession: S76082
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeduchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 B.

A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S76082
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-714 <KAN>
 A:Cross-references: EMBL:D63999; GB:AB001339; NID:G1001396; PIDN:BA010060.1; PID:G101071
 A>Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Genetics:
 A:Start codon: GTG
 C:Superfamily: Nostric sp. hypothetical protein all2707; dnaE amino-terminal homology
 F:6-70/Domain: dnaE amino-terminal homology <DNJ>

Query Match 41.8%; Score 46; DB 2; Length 714;

Best Local Similarity 53.3%; Pred. No. 28;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 2 YFPNKGTOOYTDOIE 16
 |||:~|:~|:~|:~|:
 DB 383 YFPNPSVOOYLBQLR 397

RESULT 19

T45764

A:Species: *Arabidopsis thaliana*
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 20-Jun-2000
 C:Accession: T45764
 R:Vitale, D.; Lignori, R.; Flores, M.; Arizono, A.; De Simone, V.; Mewes, H.W.; Lemcke,
 submitted to the Protein Sequence Database, December 1999
 A:Reference number: Z23012
 A:Accession: T45764
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-519 <VIT>
 A:Cross-references: EMBL:AL132980
 A:Experimental source: cultivar Columbia, BAC clone F24M12
 C:Genetics:
 A:Map position: 3
 A:Introns: 100/2; 147/3; 226/2; 303/2; 342/3; 367/2; 396/3; 419/1; 443/1
 A>Note: F24M12.370
 C:Superfamily: *Arabidopsis thaliana* hypothetical protein F24M12.380

Query Match 41.4%; Score 45.5; DB 2; Length 519;
 Best Local Similarity 55.6%; Pred. No. 24;
 Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 3 FPNKGTOOYTDOIERPIM 20
 |||:~|:~|:~|:~|:
 DB 284 FGDKG--YTDQMETPPL 298

RESULT 20

AG3278

A:Species: *Brucella melitensis*
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C:Accession: AG3278
 R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 .; Marut, M.; Coleman, E.; Selkov, E.; Elzer, P.H.; Hagins, S.; O'Callaghan, D.; Leless
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AG3278
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-299 <KUR>
 A:Cross-references: GB:AB008917; PIDN:AAL51394.1; PID:G17982098; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI0212
 A:Map position: 1

Query Match 40.9%; Score 45; DB 2; Length 299;
 Best Local Similarity 57.1%; Pred. No. 16;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 YFPNKGTOOYTDOIE 15
 |||:~|:~|:~|:~|:
 DB 162 YFPNKGTPGYDRI 175

RESULT 21

T06827

A:Species: *Pisum sativum* (garden pea)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-Jun-1999

Db 480 NAGTEQYGDMLTEMGIL 495

RESULT 26

B97208

pyruvate ferredoxin oxidoreductase CAC2499 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: B97208
R:Noiling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daily, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J: Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B97208
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1173 <KUR>
A:Cross-references: GB:AB001437; PIDN:AAK80453.1; PID:g15025520; GSPDB:GN00168
Experimental source: Clostridium acetobutylicum ATCC824
Genetics:
Gene: CAC2499

C:Superfamily: pyruvate (ferredoxin) dehydrogenase; ferredoxin 2[4Fe-4S] homology

Query Match 40.0%; Score 44; DB 2; Length 1173;
Best Local Similarity 44.4%; Pred. No. 1e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 YFPNKGTOQYTDQIERPL 19
Db 627 YQPLKGLPDPVQRIERPM 644

RESULT 27

JC4898

Down-syndrome-critical-region protein - human
N:Alternate names: Drosophila mindbrain protein homolog
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jul-2000
C:Accession: JC4898
R:Shindoh, N.; Kudoh, J.; Maeda, H.; Yamaki, A.; Minoshima, S.; Shimizu, Y.; Shimizu, N.
Biochem. Biophys. Res. Commun. 225, 92-99, 1996
A:Title: Cloning of a human homolog of the Drosophila mindbrain/rat dyrk gene from "the
A:Reference number: JC4898; MUID:96332410; PMID:8769099
A:Accession: JC4898
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-754 <SHI>
Cross-references: DDBJ:DB5759; NID:g1526445; PIDN:BA12866.1; PID:g1526446
Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 39.5%; Score 43.5; DB 2; Length 754;
Best Local Similarity 47.6%; Pred. No. 77;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

Qy 4 PNMKGTOQ-----YTDQIERPL 19
Db 46 PMSDQVSAISYSDQIQPL 66

RESULT 28

168530

hemoglobin alpha chain (clone alphaG-28) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 19-May-2000
C:Accession: 168530
R:Lam, V.M.; Gu, Y.L.; Au, D.M.; Wong, W.M.; Ma, C.W.; Cheng, L.Y.
Hemoglobin 17, 363-371, 1993
A:Title: Two new rat alpha-globin sequences as identified by the conserved region PCR.
A:Reference number: 154239; MUID:94042225; PMID:8226096
A:Accession: 168530
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-89 <RES>
A:Cross-references: GB:S6658; NID:g439804

C:Genetics:

A:Introns: 32/2

C:Superfamily: globin; globin homology

C:Keywords: blood; oxygen carrier

F:59/Binding site: oxygen (His) (distal axial ligand) #status predicted

Query Match 39.1%; Score 43; DB 2; Length 89;
Best Local Similarity 43.8%; Pred. No. 8.6;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIE 16
Db 42 TYFPNKGTOQYTDQIE 57

RESULT 29

AB0069

probable exported protein YP00563 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AB0069
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
H., M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0069
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-177 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC89419.1; PID:g15978655; GSPDB:GN00175
C:Genetics:
A:Gene: YP00563
C:Superfamily: Escherichia coli yaca protein

Query Match 39.1%; Score 43; DB 2; Length 177;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 10 QYTDQIERPLM 20
Db 144 RYTEQVERPVL 154

RESULT 30

T11169

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Smithornis sharpei mitochondrion
C:Species: mitochondrion Smithornis sharpei
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11169
R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A:Title: Multiple independent origins of mitochondrial gene order in birds.
A:Reference number: 217242
A:Accession: T11169
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-459 <MIN>
A:Cross-references: EMBL:AF090340; NID:g4894488; PID:g4894498; PIDN:AA032523.1

C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 39.1%; Score 43; DB 2; Length 459;
Best Local Similarity 45.5%; Pred. No. 53;
Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

Qy 1 TYFPNKGTOQY--TDQIERPLM 20

Db 45 TLYPYKNTLQWGTQDQISSPL 66

RESULT 31

T15851
hypothetical protein C56C10.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C:Accession: T15851

R:Fullon, L.

Submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid C56C10.

A:Reference number: Z18417

A:Accession: T15851

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1577 <FULL>

A:Cross-references: EMBL:U29468; NID:9868238; PID:9868249; PIDN:AAA68779.1; CESP:C56C10.

A:Experimental source: strain Bristol N2

A:Genetics:

A:Gene: CESP:C56C10.11

A:Introns: 40/3; 124/2; 443/1; 469/1; 521/1; 660/2; 752/3; 813/2; 868/3; 984/2; 1086/3;

C:Superfamily: dnaJ amino-terminal homology

F:786-850/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 39.1%; Score 43; DB 2; Length 1577;

Best Local Similarity 64.3%; Pred. No. 2.1e+02;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 7 GTQGYTDQIERPLM 20

Db 240 GTQHYVDIERQLV 253

RESULT 32

A44929

protein-tyrosine-phosphatase (EC 3.1.3.48) BPTP-1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Mar-1998

C:Accession: A44929

R:Adachi, M.; Sekiya, M.; Arimura, Y.; Takakawa, M.; Itoh, F.; Hinoda, Y.; Imai, K.; Yao

Cancer Res. 52, 737-740, 1992

A:Title: Protein-tyrosine phosphatase expression in pre-B cell NALM-6.

A:Reference number: A44929; NID:92119637; PMID:1370651

A:Accession: A44929

A:Molecule type: mRNA

A:Residues: 1-88 <ADA>

A:Experimental source: pre-B cell NALM-6

A:Note: sequence extracted from NCBI backbone (NCBI:78080)

A:Note: the authors translated the codon GGC for residue 12 as Gln

A:Note: the authors did not report the entire codon for residue 88

C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;

Ogy

C:Keywords: phosphoric monoester hydrolase; tyrosine-specific phosphatase

F:1-88/Domain: leukocyte common antigen cytosolic domain homology (fragment) <LAC>

Query Match 38.2%; Score 42; DB 2; Length 88;

Best Local Similarity 60.0%; Pred. No. 12;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 2 YFPNKGTOOY 11

Db 2 YFPMKCTERY 11

RESULT 33

D88970

protein F15B11.11 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: D88970

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; NID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: D88970

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <STO>

A:Cross-references: GB:chr_V; PIDN:AAB94207.1; PID:92736427; GSPDB:GN00023; CESP:F15B11.

C:Genetics:

A:Gene: F15B11.11

A:Map position: 5

Query Match 38.2%; Score 42; DB 2; Length 333;

Best Local Similarity 56.2%; Pred. No. 54;

Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 3 FPNKGTQOYDQIERP 18

Db 62 FPNKGTALDQIERP 77

RESULT 34

I56540

protein-tyrosine-phosphatase (EC 3.1.3.48) ctpl1 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 23-Jul-1999

C:Accession: I56540

R:Sahin, M.; Hockfield, S.

J. Neurosci. 13, 4968-4978, 1993

A:Title: Protein tyrosine phosphatases expressed in the developing rat brain.

A:Reference number: I56540; NID:94045925; PMID:8229209

A:Accession: I56540

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-398 <RES>

A:Cross-references: EMBL:U03273; NID:9414996; PIDN:AAC52124.1; PID:9414997

C:Genetics:

A:Gene: ctpl1

C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;

Ogy

C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase

F:1-157/Domain: protein-tyrosine-phosphatase homology (fragment) <PTP1>

F:225-398/Domain: protein-tyrosine-phosphatase homology (fragment) <PTP>

F:109/Active site: Cys (phosphocysteine intermediate) #status predicted

F:115/Binding site: substrate phosphate (Arg) #status predicted

Query Match 38.2%; Score 42; DB 2; Length 398;

Best Local Similarity 60.0%; Pred. No. 66;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 2 YFPNKGTOOY 11

Db 24 YFPMKCTERY 33

RESULT 35

B86164

hypothetical protein F15K9.13 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C:Accession: B86164

R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

```

A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: B86164
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-514 <STO>
A:Cross-references: GB:AB005172; NID:G3850587; PIDN:AACT2127.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match      38.2%  Score 42; DB 2; Length 514;
Best Local Similarity 46.7%  Pred. No. 88;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      6  KGTQGYTDQIERPLM 20
Db      495  RGTNNTTKITEPPL 509

RESULT 36
A:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: AB5075
R:Anonymous, The European Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: AB5001; MUID:20083488; PMID:10617196
A:Cross-references: GB:NC_001268; NID:G7267360; PIDN:CAB81133.1; GSPDB:GN00140
C:Genetics:
A:Gene: ATG907630
A:Map position: 4

Query Match      38.2%  Score 42; DB 2; Length 538;
Best Local Similarity 47.1%  Pred. No. 92;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      1  TYPNKGTYQYTDQIER 17
Db      74  TSYPHKGTQIEREKER 90

RESULT 37
A:Species: Plasmodium falciparum
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jun-2000
C:Accession: S14535
R:Schreiber, L.; Deutsche, U.; Stork, T.; Mueller-Hill, D.
submitted to the EMBL Data Library, December 1989
A:Reference number: S14469
A:Accession: S14535
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-669 <SCH>
A:Cross-references: EMBL:X17486; NID:G9848; PID:G9849

Query Match      38.2%  Score 42; DB 2; Length 669;
Best Local Similarity 47.4%  Pred. No. 12e+02;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY      1  TYPNKGTYQYTDQIERPL 19
Db      222  THFNNNLINQYVDQNIPL 240

RESULT 38
A:Species: beta-glucosidase (EC 3.2.1.21) - Salmonella enterica subsp. enterica serovar Typhi (str

```

```

C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AD0778
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AD0778
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-765 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02546.1; PID:G16503407; GSPDB:GN00176
C:Genetics:
A:Gene: STY2396
C:Superfamily: beta-glucosidase
C:Keywords: glycosidase; hydrolase

Query Match      38.2%  Score 42; DB 2; Length 765;
Best Local Similarity 50.0%  Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      4  PNKGTQYTDQIERPL 19
Db      624  PNKTSRYFDANGPL 639

RESULT 39
A:Species: Glycine max (soybean)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T06429
R:Bunker, T.W.; Koetje, D.S.; Stephenson, L.C.; Creelman, R.A.; Mullet, J.E.; Grimes, H.
Plant Cell 7, 1319-1331, 1995
A:Title: Sink limitation induces the expression of multiple soybean vegetative lipoxigen
A:Reference number: Z15673; MUID:96004535; PMID:7549487
A:Accession: T06429
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-859 <BDN>
A:Cross-references: EMBL:U26457; NID:G1262439; PIDN:AAA96817.1; PID:G1262440
A:Experimental source: cultivar Wye; leaf
C:Genetics:
A:Gene: V1XC
C:Function:
A:Description: catalyzes hydroperoxidation of polyunsaturated fatty acids containing a c
C:Keywords: iron; metalloprotein; oxidoreductase

Query Match      38.2%  Score 42; DB 2; Length 859;
Best Local Similarity 29.4%  Pred. No. 1.5e+02;
Matches 5; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY      2  YPNKGTYQYTDQIERP 18
Db      735  FIPERGTXDEBMDP 751

RESULT 40
A:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: B48758
R:Pan, M.G.; Rim, C.; Lu, K.P.; Florio, T.; Stork, P.J.S.
U. Biol. Chem. 268, 19284-19291, 1993
A:Title: Cloning and expression of two structurally distinct receptor-linked protein-tyr
A:Reference number: B48758; MUID:93374907; PMID:8396131
A:Accession: B48758

```

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1262 <PAN>
 A:Cross-references: GB:L19181
 A:Note: the authors translated the codon TGC for residue 27 as Gly, GAG for residue 79 a
 727 as Phe; residues 1237-1244 are shown after residue 1262, and, consequently, residues
 C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology/
 ogy
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; trans
 F:149-209/Domain: immunoglobulin homology <IMM1>
 F:246-300/Domain: immunoglobulin homology <IMM2>
 F:318-398/Domain: fibronectin type III repeat homology <FN3A>
 F:411-497/Domain: fibronectin type III repeat homology <FN3B>
 F:509-590/Domain: fibronectin type III repeat homology <FN3C>
 F:600-677/Domain: fibronectin type III repeat homology <FN3D>
 F:880-1262/Domain: leukocyte common antigen cytosolic domain homology #status atypical
 F:967-1188/Domain: protein-tyrosine-phosphatase homology <PTP1>
 F:1140/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1146/Binding site: substrate phosphate (Arg) #status predicted
 Query Match 38.2%; Score 42; DB 1; Length 1262;
 Best Local Similarity 60.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YFPNKGTOOY 11
 |:|:|:|:|
 Db 1055 YWPNRGTEY 1064

Search completed: August 20, 2003, 09:32:45
 Job time : 23.4742 secs


```

DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
CN MTND4 OR ND4 OR NADH4.
OS Struthio camelus (Oestrich).
OC Mitochondrion.
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Palaeognathae; Struthioniformes; Struthionidae;
CX NCBI_Taxid=8801;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=97357422; PubMed=9214748;
RA Harlid A., Janke A., Arnason U.;
RT "The mtDNA sequence of the ostrich and the divergence between
RL palaeognathous and neoognathous birds".
RN Mol. Biol. Evol. 14:754-761(1997).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21263106; PubMed=11370967;
RA Haddrath O., Baker A.J.;
RT "Complete mitochondrial DNA genome sequences of extinct birds: ratite
phylogenetics and the vicariance biogeography hypothesis.";
Proc. R. Soc. Lond., B, Biol. Sci. 268:939-945(2001).
CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Y12025; CAA72753.1; -.
DR EMBL; AF387815; AAK5354.1; -.
DR PIR; B99613; B99613.
DR PIR; T11528; T11528.
DR InterPro; IPR003918; NADhub oxxred4.
DR InterPro; IPR001750; Oxidored_g1.
DR InterPro; IPR000260; Oxidored_g5_N.
DR Pfam; PF00361; oxidored_g1; 1.
DR Pfam; PF01059; oxidored_g5_N; 1.
DR PRINTS; PR01437; NUOXDRPTASB4.
DR Oxidoreductases; NAD; Ubiquinone; Mitochondrion.
KM SEQUENCE 458 AA; 50955 MW; 575118B6297A18AC CRC64;
SQ
Query Match 46.4%; Score 51; DB 1; Length 458;
Best Local Similarity 50.0%; Pred. No. 0.83;
Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 1;
DB 45 TYFPKNTGTPWTGIDISAPLL 66
1 TYFPNKGTQOYTT--DQIERPLM 20
|||||:|||||:
|||:|:|:|:|:|:
1 45 TYFPKNTGTPWTGIDISAPLL 66
RESULT 3
MGA_HUMAN STANDARD; PRT; 1856 AA.
AC O43451;
DT 30-MAY-2000 (Rel. 39, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DS Maltase-glucoamylase, intestinal [Includes: Maltase (EC 3.2.1.20)
DS (Alpha-glucosidase), Glucoamylase (EC 3.2.1.3) (Glucan 1,4-alpha-
DS glucosidase)].
GN MGAM OR MGA OR MGAML.
GS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.
NCBI_Taxid=9606;
[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Small intestine;
RX MEDLINE=98112863; PubMed=9446624;

```

RA	Nichols B.L., Eldering J.A., Avery S.E., Hahn D., Quaroni A.,
RA	Steerchi E.E.,
RT	"Human small intestinal maltase-glucoamylase cDNA cloning. Homology
RT	to sucrase-isomaltase.";
RT	J. Biol. Chem. 273:3076-3081 (1998).
RN	[2]
RP	REVISIONS TO 776: 1049; 1100; 1541; 1612 AND 1811.
RA	Nichols B.L., Eldering J.A., Avery S.E., Hahn D., Quaroni A.,
RA	Steerchi E.E.,
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RP	[3]
RP	CHARACTERIZATION.
RC	TISSUE=Small intestine mucosa;
RX	MEDLINE=89066802; PubMed=3143729;
RX	Nam H.Y., Steerchi E.E., Lentze M.J.;
RT	"Structure, bioynthesis, and glycosylation of human small intestinal
RT	maltase-glucoamylase.";
RL	J. Biol. Chem. 263:19709-19717 (1988).
RL	[4]
RP	SULFATION
RX	MEDLINE=88082658; PubMed=3121301;
RA	Danielson E.M.;
RT	"Tyrosine sulfation, a post-translational modification of microvillar
RT	enzymes in the small intestinal enterocyte.";
EMO	J. 6:2891-2896(1987).
CC	-1- FUNCTION: MAY SERVE AS AN ALTERNATE PATHWAY FOR STARCH DIGESTION
CC	WHEN LUMINAL ALPHA-AMYLASE ACTIVITY IS REDUCED BECAUSE OF
CC	IMMATUREITY OR MALNUTRITION. MAY PLAY A UNIQUE ROLE IN THE
CC	DIGESTION OF MALTED DIETARY OLIGOSACCHARIDES USED IN FOOD
CC	MANUFACTURING.
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-
CC	linked D-glucose residues with release of D-glucose.
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
CC	glucose residues successively from non-reducing ends of the chains
CC	with release of beta-D-glucose.
CC	-1- SUBUNIT: Monomer.
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein. Brush border.
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN SMALL INTESTINE, GRANULOCYTE, AND
CC	KIDNEY BUT NOT IN SALIVARY GLAND OR PANCREAS.
CC	-1- PTM: N- AND O-GLYCOSYLATED.
CC	-1- PTM: DOES NOT UNDERGO INTRACELLULAR OR EXTRACELLULAR PROTEOLYTIC
CC	CLEAVAGE.
CC	-1- PTM: SULFATED (BY SIMILARITY).
CC	-1- SIMILARITY: BELONGS TO FAMILY 31 OF GLYCOSYL HYDROLASES.
CC	-1- SIMILARITY: Contains 2 P-type (trefoil) domains.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb.ch/announce/
CC	or send an email to license@isb-ib.ch).
CC	-----
DR	EMBL; AF016833; AAC39568.2; -.
DR	Genew; HGNC:7043; MGAM.
DR	MI; 154360; -.
DR	GO; GO:0005963; P:starch catabolism; TAS.
DR	InterPro; IPR000322; Glyco_hydro_31.
DR	InterPro; IPR000519; P_trefoil1.
DR	Pfam; PF01055; Glyco_hydro_31; 2.
DR	Pfam; PF00088; trefoil1; 2.
DR	SMART; SMO0018; pd; 2.
DR	PROSITE; PS00129; GLYCOSYL_HYDROL_F31_1; 2.
DR	PROSITE; PS00025; P_TREFOIL; 1.
KM	MultiFunctional enzyme; Transmembrane; Glycoprotein; Hydrolase;
KM	Glycosidase; Repeat; Signal-anchor; Sulfation.
FT	INT MET 0
FT	DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 13 33 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT	DOMAIN 34 1856 (POTENTIAL).
FT	DOMAIN 37 83 LUMENAL (POTENTIAL).
FT	SER/THR-RICH.


```

FT DOMAIN 88 132 P-TYPE 1.
FT DOMAIN 197 914 MALTASE.
FT DOMAIN 952 998 P-TYPE 2.
FT DOMAIN 1066 1812 GLUCOMALYASE.
FT ACT_SITE 528 528 BY SIMILARITY.
FT ACT_SITE 1419 1419 BY SIMILARITY.
FT DISULFID 89 117 BY SIMILARITY.
FT DISULFID 100 116 BY SIMILARITY.
FT DISULFID 111 129 BY SIMILARITY.
FT DISULFID 965 982 BY SIMILARITY.
FT DISULFID 977 995 BY SIMILARITY.
FT MOD_RES 415 415 SULFATION (POTENTIAL).
FT MOD_RES 424 424 SULFATION (POTENTIAL).
FT MOD_RES 1281 1281 SULFATION (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 706 706 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 826 826 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 884 884 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 976 976 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 988 988 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1254 1254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1332 1332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1363 1363 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1387 1387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1602 1602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1671 1671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1841 1841 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1846 1846 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1856 AA; 209721 MW; 81E7A0CA8ABD07D CRC64;

```

```

Query Match 44.5%; Score 49; DB 1; Length 1856;
Best Local Similarity 58.8%; Pred. No. 8.5;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 PNRKQOYTDQIERPLM 20
DB 1221 PELVYQYTELGRPVM 1237

```

```

RESULT 4
SEC3_YEAST STANDARD; PRT; 1336 AA.
P33332;

```

```

DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Exocyst complex component SEC3 (PSLI protein).
GN SEC3 OR PSL1 OR YER008C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9704444; PubMed=8889515;
RA Haarer B.K., Corbett A., Kwon Y., Petzold A.S., Silver P.,
RA Brown S.S.;
RT "SEC3 mutations are synthetically lethal with profilin mutations and
RT cause defects in diploid-specific bud-site selection.";
RL Genetics 144:495-510(1996).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX PubMed=9169868;
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA Arayo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,

```

```

RA Humicke-Smith S., Hyman R.W., Kayser A., Komp C., Laakkari D., Lew H.,
RA Lin D., Moseedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL Nature 387:78-81(1997).

```

```

[3]
CHARACTERIZATION.
RP MEDLINE=97390810; PubMed=9247645;
RA Finger F.P., Novick P.;
RT "Sec3p is involved in secretion and morphogenesis in Saccharomyces
RT cerevisiae.";
RL Mol. Biol. Cell 8:647-662(1997).

```

```

[4]
SEQUENCE OF 793-800.
RP MEDLINE=97133278; PubMed=8978675;
RX Terbush D.R., Maurice T., Roth D., Novick P.;
RT "The Exocyst is a multi-protein complex required for exocytosis in
RT Saccharomyces cerevisiae.";
RL EMO J. 15:6483-6494(1996).

```

```

CC -1- FUNCTION: Component of the exocyst complex involved in the docking
CC of exocytic vesicles with fusions site on the plasma membrane.
CC -1- SUBUNIT: The exocyst complex is composed of SEC3, SEC5, SEC6,
CC SEC8, SEC10, SEC15, EXO70 and EXO84.
CC -1- SIMILARITY: BELONGS TO THE SEC3 FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; L22204; AAB49380.1; -
DR EMBL; U18778; AAB64541.1; -
DR PIR; S41794; S41794.
DR SGD; S0000810; SEC3.
DR GO; GO:0000143; C:actin cap (sensu Saccharomycos); IDA.
DR GO; GO:0016288; P:cytokinesis; IMP.
DR GO; GO:0007121; P:polar budding; IMP.
DR GO; GO:0007121; P:polar budding; IMP.
DR EXOCYSTOSIS; Transport; Protein transport; Coiled coil.
FT DOMAIN 1309 1336 COILED COIL (POTENTIAL).
FT DOMAIN 1309 1336 COILED COIL (POTENTIAL).
SQ SEQUENCE 1336 AA; 154694 MW; 6F0C1BF13A77985E CRC64;

```

```

Query Match 43.6%; Score 48; DB 1; Length 1336;
Best Local Similarity 42.1%; Pred. No. 8.6;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 TYPRKQOYTDQIERPL 19
DB 498 TYPRSTARAHDDQKPL 516

```

```

RESULT 5
VGLG_IHNV STANDARD; PRT; 508 AA.
ID_VGLG_IHNV
AC P07923;

```

```

DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Spike glycoprotein precursor.
GN G.
OS Infectious hematopoietic necrosis virus (strain Round Butte) (IHNV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Novirhabdovirus.
OC NCBI_TaxID=11291;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87198856; PubMed=3033264;
RA Koener J.F., Passavant C.W., Kurath G., Leong J.;
RA "Nucleotide sequence of a cDNA clone carrying the glycoprotein gene

```

RT of infectious hematopoietic necrosis virus, a fish rhabdovirus.";
 RL J. Virol. 61:1342-1349(1987).
 RN [2]
 RP REVISIONS.
 RA Leong J.;
 RL Submitted (XXX-1995) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M16023; AAB59926.1; -
 DR PIR; A29532; VGVNRR. -
 DR InterPro; IPR001903; Rhabd_glycop.
 DR InterPro; IPR002417; Spike_prot.
 DR Pfam; PF00974; Rhabdo_glycop; 1.
 DR PRINTS; PR00796; SPIKEPROTEIN.
 DR Transmembrane; Envelope protein; Glycoprotein; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 508 SPIKE GLYCOPROTEIN. (POTENTIAL).
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 506 506 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 508 AA; 56799 MW; CF65376C626F89B CRC64;
 Query Match 40.0%; Score 44; DB 1; Length 508;
 Best Local Similarity 42.1%; Pred. No. 14;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 2 YFPNKGTOQYTDQIERPLM 20
 DB 144 YFPAPRCQWYTDVQNDLI 162
 RESULT 5
 CH60_PASMU STANDARD; PRT; 547 AA.
 ID CH60_PASMU
 AC 059687;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DT 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 GN GROEL OR GROEL OR MOXA OR PM107.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=P1059;
 RA MEDLINE=96105224; PubMed=8529887;
 RA Love B.C., Hansen L.W., Hirsch D.C.;
 RT "Cloning and sequence of the groESL heat-shock operon of Pasteurella
 RT multocida.";
 RL Gene 166:179-180(1995).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RA MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
 CC proper assembly of unfolded polypeptides generated under stress
 CC conditions (By similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
 CC 7 subunits (By similarity).
 CC -1- SIMILARITY: Oligomer of 14 subunits composed of two stacked rings of
 CC 7 subunits (By similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U30165; AAB4916.1; -
 DR EMBL; AB006151; AAA03191.1; -
 DR PIR; JC4519; JC4519.
 DR HSSP; P06139; JCON.
 DR HAMAP; MF_00600; -; 1.
 DR InterPro; IPR001844; Chaperin Cpn60.
 DR InterPro; IPR002423; Cpn60/TCF-1.
 DR Pfam; PF00118; Cpn60_TCF1; 1.
 DR PRINTS; PR00298; CHAPERONING6.
 DR PRINTS; PR00304; TCOMPLEXTCF1.
 DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
 DR Chaperone; ATP-binding; Complete proteome.
 FT CONFLICT 424 424 S -> N (IN REF. 1).
 FT CONFLICT 464 464 I -> V (IN REF. 1).
 SQ SEQUENCE 547 AA; 57291 MW; EBB95B9F235B855E CRC64;
 Query Match 40.0%; Score 44; DB 1; Length 547;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 5 NKGTOQYTDQIERPLM 20
 DB 479 NAGTQYGDMEIMGL 494
 RESULT 7
 CH60_HABIN STANDARD; PRT; 548 AA.
 ID CH60_HABIN
 AC P43733;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DT 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 GN GROEL OR MOXA OR GROEL OR H10543.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RA MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fitchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
 CC proper assembly of unfolded polypeptides generated under stress
 CC conditions (By similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
 CC 7 subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U32736; AAC2201.1; -
 DR PIR: G64076; G64076.
 DR HSSP: P06139; IJON.
 DR TIGR: H10543; -
 DR HAMAP: MF_00600; -; 1.
 DR InterPro: IPR001844; Chaperin Cpn60.
 DR InterPro: IPR002423; Cpn60/TCF-1.
 DR Pfam: PF00118; Cpn60_TCF1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 DR Chaperone; ATP-binding; Complete proteome.
 DR SEQUENCE 548 AA; 57577 MW; CA4066AAC1B62159 CRC64;

Query Match 40.0%; Score 44; DB 1; Length 548;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTOYTDOTERPLM 20
 DB 480 NAGTEQYDWMLEMTL 495

RESULT 8
 DYRA_HUMAN STANDARD; PRT: 763 AA.
 AC Q13627; Q60769; Q92582; Q92810; Q9UNM5;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Dual-specificity tyrosine-phosphorylation regulated kinase 1A
 DE (EC 2.7.1.-) (Protein kinase minibrain homolog) (MNBH) (HP86) (Dual
 DE specificity YAKI-related kinase).
 GN DYR1A OR DYRK OR MNBH OR MNB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RX MEDLINE=97131512; PubMed=8975710;
 RA Song W.J., Sternberg L.R., Kaestlen-Sportes C., van Keuren M.L.,
 RA Chung S.H., Slack A.C., Miller D.E., Glover T.W., Chiang P.W., Lou L.,
 RA Kunitz D.W.;
 RA "Isolation of human and murine homologues of the Drosophila minibrain
 RA gene: human homologue maps to 21q22.2 in the Down syndrome 'critical
 RA region.'";
 RT Genomics 38:331-339(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RX MEDLINE=97026291; PubMed=8872470;
 RA Guimera J., Casas C., Pucharcos C., Solans A., Domenech A.,
 RA Planas A.M., Ashley J., Lovett M., Estivill X., Pritchard M.A.;
 RA "A human homologue of Drosophila minibrain (MNB) is expressed in the
 RA neuronal regions affected in Down syndrome and maps to the critical
 RA region.";
 RT Hum. Mol. Genet. 5:1305-1310(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX TISSUE=Fetal brain;
 RX MEDLINE=96332410; PubMed=8769099;
 RA Shindzu N., Kudoh J., Maeda H., Yamaki A., Minoshima S., Shimizu Y.,
 RA Shindzu N.;
 RA "Cloning of a human homolog of the Drosophila minibrain/rat Dyrc gene
 RA from 'the Down syndrome critical region' of chromosome 21.";

RL Biochem. Biophys. Res. Commun. 225:92-99(1996).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Miki O., Masahira H., Yasufumi M., Naohiko S., Takahiro N.,
 RA Hitoshi I., Nobuo N., Toshiko E., Yoshiyuki S., Misa O.,
 RA Ohtera M., Seki N., Nagase T., Suzuki E., Nomura N., Ohara O.,
 RA Hattori M., Sakaki Y., Eki T., Murakami Y., Saito T., Ichikawa H.,
 RA Ohts M.;
 RA Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A., VARIANTS PHE-415 AND HIS-681, AND ALTERNATIVE
 RP SPLICING.
 RX MEDLINE=99263503; PubMed=10329007;
 RA Guimera J., Casas C., Estivill X., Pritchard M.;
 RA "Human minibrain homologue (MNBH/DYRK1): characterization, alternative
 RA splicing, differential tissue expression, and overexpression in Down
 RA syndrome.";
 RT Genomics 57:407-418(1999).
 RN [6]
 RP SEQUENCE OF 234-380 FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=98163741; PubMed=9503011;
 RA Dahmane N., Alt Ghezala G., Gossel P., Chamoun Z.,
 RA Dutresne-Zacharia M.C., Lopes C., Rabatel N., Gassanova-Maugenre S.,
 RA Chetouh Z., Abramowski V., Fayet E., Yasp M.L., Korn B.,
 RA Blouin J.L., Lehnach H., Pousetka A., Antonarakis S.E., Sinet P.M.,
 RA Creau N., Delabar J.M.;
 RA "Transcriptional map of the 2.5-Mb CBR-ERG region of chromosome 21
 RA involved in Down syndrome.";
 RT Genomics 48:12-23(1998).
 CC -1- FUNCTION: MAY PLAY A ROLE IN A SIGNALING PATHWAY REGULATING
 CC NUCLEAR FUNCTIONS OF CELL PROLIFERATION. PHOSPHORYLATES SERINES,
 CC THREONINES AND TYROSINES RESIDUES IN ITS SEQUENCE AND IN EXOGENOUS
 CC SUBSTRATES.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=5;
 CC Comment=Additional isoforms seem to exist;
 CC Name=long; Sequence=displayed;
 CC IsoId=Q13627-1; Sequence=VSP_004917;
 CC Name=1;
 CC IsoId=Q13627-2; Sequence=VSP_004917;
 CC Name=2;
 CC IsoId=Q13627-3; Sequence=VSP_004918, VSP_004919;
 CC Name=3;
 CC IsoId=Q13627-4; Sequence=VSP_004920, VSP_004921;
 CC Name=4;
 CC IsoId=Q13627-5; Sequence=VSP_004922, VSP_004923;
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS. HIGHEST LEVELS IN SKELETAL MUSCLE,
 CC TESTIS, FETAL LUNG AND FETAL KIDNEY.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE DEVELOPING CENTRAL NERVOUS
 CC SYSTEM.
 CC -1- PTM: AUTOPHOSPHORYLATED ON TYR RESIDUES.
 CC -1- DISEASE: OVEREXPRESSED 1.5-FOLD IN FETAL DOWN SYNDROME BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MNB/DYRK SUBFAMILY

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U58496; AAC50939.1; -
 DR EMBL: U52173; AAB18639.1; -
 DR EMBL: D85759; AAB12866.1; -
 DR EMBL: D86550; BAA13110.1; -
 DR EMBL: AF108830; AAD11169.1; -
 DR EMBL: AJ001870; CAA05059.1; -
 DR PIR: JC4898; JC4898.
 DR HSSP: Q00534; 1B18.

ID	Category	Value	Description
DR	GeneW; HGNC:3091; DYRK1A.		
DR	MIM; 600855; -		
DR	GO; GO:0005634; C:nucleus; TAS.		
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.		
DR	GO; GO:0007359; P:neurogenesis; TAS.		
DR	InterPro; IPR000719; Prot kinase.		
DR	InterPro; IPR002290; Ser Thr_kinase.		
DR	Pfam; PF00069; pkinase_1		
DR	PRODOM; PD000001; Prot_kinase; 1.		
DR	SMART; SMO0220; S_TKc_1.		
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.		
DR	PROSITE; PS00108; PROTEIN KINASE SF; 1.		
DR	PROSITE; PSS0011; PROTEIN KINASE DOM; 1.		
KM	Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase		
KM	ATP-binding; Nuclear protein; Phosphorylation; Alternative splicing;		
KM	Polymorphism.		
FT	DOMAIN	117	134
FT	NP_BIND	159	479
FT	BINDING	165	173
FT	ACT_SITE	188	186
FT	DOMAIN	287	287
FT	DOMAIN	509	515
FT	DOMAIN	599	602
FT	DOMAIN	607	619
FT	DOMAIN	656	672
FT	DOMAIN	664	671
FT	MOD_RES	219	219
FT	MOD_RES	319	319
FT	MOD_RES	321	321
FT	VASPLIC	70	78
FT	VASPLIC	516	529
FT	VASPLIC	530	763
FT	VASPLIC	516	540
FT	VASPLIC	541	763
FT	VASPLIC	559	584
FT	VASPLIC	585	763
FT	VARIANT	415	415
FT	VARIANT	681	681
FT	CONFLICT	32	32
FT	CONFLICT	47	47
FT	CONFLICT	57	57
FT	CONFLICT	123	123
FT	CONFLICT	266	266
FT	CONFLICT	357	357
FT	CONFLICT	397	397
FT	CONFLICT	592	592
FT	CONFLICT	679	679
FO	SEQUENCE	763 AA; 85584 MM; 7C3A52A3CB8D4FB5 CRC64;	

Query Match	39.5%;	Score 43.5;	DB 1;	Length 763;
Best Local Similarity	47.6%;	Pred. No. 26;		
Matches 10;	Conservative 3;	Mismatches 3;	Indels 5;	Gaps 1

QY 4 PNKGTQ-----YTDQIERPL 19
||| |::|||
Db 46 PNISDQVSALSYSDQIQPL 66

RESULT 9
DYRA_RAT

ID	DRYA_RAT	STANDARD;	PRT;	763 AA.
AD	Q63470;			
DT	15-JUL-1999	(Rel. 38, Created)		
DT	15-JUL-1999	(Rel. 38, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Dual-specificity tyrosine-phosphorylation regulated kinase 1A			
DE	(EC 2.7.1.-) (Protein kinase minibrain homolog) (MNBH) (RP86) (Dual			
DE	specificity YAK1-related kinase).			
GN	DYRK1A OR DYRK.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_Taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF TYR-219.			
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;			
RC	MEDLINE=96216443; PubMed=6631952;			
RA	Kentrup H., Becker W., Henkelbach J., Wilmes A., Schuermann A.,			
RA	Huppertz C., Kainulainen H., Joost H.-G.;			
RT	"DYRK, a dual specificity protein kinase with unique structural			
RT	features whose activity is dependent on tyrosine residues between			
RT	subdomains VII and VIII."			
RL	J. Biol. Chem. 271:3488-3495(1996).			
RN	[2]			
RP	REVISIONS.			
RA	Kentrup H.;			
RL	Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	CHARACTERIZATION.			
RX	MEDLINE=98421512; PubMed=9748265;			
RA	Becker W., Weber Y., Wetzel K., Bimblter K., Tejedor F.J.,			
RA	Joost H.-G.;			
RT	"Sequence characteristics, subcellular localisation, and substrate			
RT	specificity of DYRK-related kinases, a novel family of dual			
RT	specificity protein kinases."			
RL	J. Biol. Chem. 273:25893-25902(1998).			
CC	-1- FUNCTION. MAY PLAY A ROLE IN A SIGNALING PATHWAY REGULATING			
CC	NUCLEAR FUNCTIONS OF CELL PROLIFERATION. PHOSPHORYLATES SERINES,			
CC	THREONINES AND TYROSINES RESIDUES IN ITS SEQUENCE AND IN EXOGENOUS			
CC	SUBSTRATES.			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event-Alternative splicing; Named isoforms=2;			
CC	Name=Long;			
CC	IsoId=Q63470-1; Sequence=Displayed;			
CC	Name=Short;			
CC	IsoId=Q63470-2; Sequence=VSP_004924;			
CC	-1- TISSUE SPECIFICITY: Ubiquitous.			
CC	-1- PTM: AUTOPHOSPHORYLATED ON TYR RESIDUES (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	MNB/DYRK SUBFAMILY.			
CC	-----			
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION			
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION			
CC	IN THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS			
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY			
CC	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL			
CC	ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sib.ch/announce/			
CC	OR SEND AN EMAIL TO license@isb-sib.ch).			
CC	-----			
DR	EMBL; X79769; CA56164.1; -.			
DR	HSSP; Q00534; 1B18.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR002290; Ser_thr_pkinase.			
DR	Pfam; PF00069; pkinase; 1.			
DR	ProDom; PD000001; Prot_kinase; 1.			
DR	SMART; SM0220; S_TKc; 1.			
DR	PROSITE; PS00107; PROTEIN KINASE_ATP_1.			
DR	PROSITE; PS00108; PROTEIN KINASE_ST_1.			
DR	PROSITE; PS50011; PROTEIN KINASE_DOM_1.			
KW	Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;			
KW	ATP-binding; Nuclear protein; Phosphorylation; Alternative splicing.			
FT	DOMAIN	117	134	BIPARTITE NUCLEAR LOCALIZATION SIGNAL

```

FT DOMAIN 159 479 (POTENTIAL).
FT NP BIND 165 173 PROTEIN KINASE.
FT BINDING 188 188 ATP (BY SIMILARITY).
FT ACT SITE 287 287 ATP (BY SIMILARITY).
FT DOMAIN 509 515 BY SIMILARITY.
FT DOMAIN 599 602 POLY-SER.
FT DOMAIN 607 619 POLY-HIS.
FT DOMAIN 656 672 POLY-HIS.
FT MOD_RES 219 219 SER/THR-RICH.
FT MOD_RES 319 319 PHOSPHORYLATION (AUTO-).
FT MOD_RES 321 321 PHOSPHORYLATION (AUTO-).
FT VARSPLIC 70 78 Missing (in isoform Short).
FT MUTAGEN 219 219 /FTid=VSP 004924.
FT Y->F: REDUCED AUTOPHOSPHORYLATION ON
FT TYROSINE, BUT NO LOSS OF HISTONE
FT PHOSPHORYLATION.
SQ SEQUENCE 763 AA; 85541 MW; CBSECTECACIF9A47 CRC64;

Query Match
Best Local Similarity 47.6%; Score 43.5; DB 1; Length 763;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

Oy 4 PNRKGTQO-----YTDQIERPL 19
Db 46 PNRDQVSAISYSQIQQL 66

RESULT 10
CH60 ACTPL STANDARD; PRT; 546 AA.
AC P94166;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
OS GROEL OR GROEL OR MOXA.
OC Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5 4074 / Serotype 1;
RX MEDIAN=97189570; PubMed=9037757;
RA Vezina G., Strole M., Claitoux N., Boissinot M.;
"Cloning and characterization of the groE locus from Actinobacillus
pleuropneumoniae.";
FEMS Microbiol. Lett. 147:11-16 (1997).
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U55016; AAB51437.1; -.
CC HSSP; P06139; IGRU.
CC HAMAP; MF_00600; -. 1.
CC InterPro; IPR001844; Chaperonin Cpn60.
CC InterPro; IPR002423; Cpn60/ICP-1.
CC Pfam; PF00118; cpn60_TCP1. 1.
CC PRINTS; PR00298; CHAPERONIN60.

```

```

DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONIN_CPN60; 1.
KW Chaperone; ATP-binding.
FT INT MET 0
SQ SEQUENCE 546 AA; 57513 MW; E84B872C9BD3DB56 CRC64;

Query Match
Best Local Similarity 42.1%; Score 43; DB 1; Length 546;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Oy 2 YPNKGTQOYTDQIERPLM 20
Db 476 YGNAGTEQYDMLMGIL 494

```

```

RESULT 11
G7AC_BREDI STANDARD; PRT; 720 AA.
ID G7AC_BREDI
AC Q915D6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glutaryl 7-aminocapthalosporanic acid acylase precursor (EC 3.5.1.-)
DE (Glutaryl-7-ACA-acylase).
OS Brevundimonas diminuta (Pseudomonas diminuta).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Brevundimonas.
OX NCBI_TaxID=293;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KAC-1;
RA Kim D.-W., Kang S.-M., Yoon K.-H.;
"Cloning and the nucleotide sequence of a Pseudomonas diminuta KAC-1
glutaryl 7-aminocapthalosporanic acid acylase gene.";
Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
CC -1- PATHWAY: SECOND STEP IN THE BIOCONVERSION OF CEPHALOSPORIN TO 7-
CC ACA.
CC -1- SUBCELLULAR LOCATION: Periplasmic (potential).
CC -1- SIMILARITY: BELONGS TO PERITRANSFAM FAMILY 845.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF251710; AAF64242.1; -.
CC PDB; 1FM2; 15-AUG-01.
CC PDB; 1TVZ; 01-SEP-02.
CC PDB; 1JW0; 01-SEP-02.
CC PDB; 1KEH; 04-DEC-02.
CC MEROPS; S45.002; -.
CC InterPro; IPR002692; Penicill amidase.
CC Pfam; PF01804; Penicill amidase; 1.
CC Hydrolase; Periplasmic; Antibiotic resistance; Zymogen; Signal;
CC 3D-structure.
CC 3D-structure.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 720 GLUTARYL 7-AMINOCEPHALOSPORANIC ACID
FT ACYLASE.
FT ACT SITE 199 199 PROBABLE.
SQ SEQUENCE 720 AA; 79779 MW; AD624797845CC39B CRC64;

Query Match
Best Local Similarity 39.1%; Score 43; DB 1; Length 720;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 7 GTQOYTDQIER 17
Db 680 GTTHYSQIER 690

```

RESULT 12
UBIQ_HUMAN STANDARD; PRT; 913 AA.
AC 09BXU7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase 26 (EC 3.1.2.15) (Ubiquitin
thioesterase 26) (Ubiquitin-specific processing protease 26)
GN USP26.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21175748; PubMed=11279525;
Wang P.J., McCarrey J.R., Yang F., Page D.C.;
"An abundance of X-linked genes expressed in spermatogonia.";
Nat. Genet. 27:422-426(2001).
CC -1- FUNCTION: INVOLVED IN THE UBIQUITIN-DEPENDENT PROTEOLYTIC PATHWAY
IN CONJUNCTION WITH THE 26S PROTEASOME (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thioester + H(2)O =
ubiquitin + a thiol.
CC -1- SIMILARITY: Belongs to peptidase family C19.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; AF285593; AAK31972.1; -
DR MEROPS; C19.046; -
DR Genew; HGNC:13485; USP26.
DR MIM; 300309; -
DR InterPro; IPR001394; UCH-2.
DR Pfam; PF00443; UCH; 1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
CC Ubiquitin conjugation pathway; Hydrolase; Thiol protease; Multigene family.
FT ACT_SITE 304 304 BY SIMILARITY.
ACT_SITE 832 832 BY SIMILARITY.
ACT_SITE 841 841 BY SIMILARITY.
SEQUENCE 913 AA; 104047 MW; BB70B09CDCBAJ348 CRC64;

Query Match 39.1%; Score 43; DB 1; Length 913;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Db 760 FRKPGTQGHTRKULRP 775

RESULT 13
UBIG_RHIME STANDARD; PRT; 248 AA.
AC 092MKL;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-demethylubiquinone-9-3-methyltransferase (EC 2.1.1.64) (3,4-
dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
methyltransferase)
GN UBIG OR R02615 OR SMC02440.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_Taxid=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
Boitard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
Godie T., Goffeau A., Kahn D., Kiss E., Lelaune V., Maury D.,
Rahat T., Portetelle D., Puehler A., Purnelle B., Rampersingh U.,
Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
"Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium meliloti strain 1021.";
Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; AL591791; CAC47194.1; -
DR HAMAP; MF_00472; -; 1.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM bind.
KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 248 AA; 27438 MW; 0190692A5C18C593 CRC64;

Query Match 38.6%; Score 42.5; DB 1; Length 248;
Best Local Similarity 47.4%; Pred. No. 11;
Matches 9; Conservative 3; Mismatches 2; Indels 5; Gaps 1;

Db 190 RGTQYDRLVPRDELRLP 208

RESULT 14
BGX_SALTY STANDARD; PRT; 765 AA.
AC 056078;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Periplasmic beta-glucosidase precursor (EC 3.2.1.21) (Gentibiase)
(Cellulase) (Beta-D-glucoside glucohydrolase) (T-cell inhibitor).
GN BGX OR STM2166.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_Taxid=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RA Matouli K.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanders R.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Portolillo S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
"Complete genome sequence of Salmonella enterica serovar Typhimurium

```

RT LT2."
RC Nature 413:852-856(2001).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC glucose residues with release of beta-D-glucose.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ibt-sib.ch/announce/
CC or send an email to license@ibt-sib.ch).
CC
CC EMBL: D86507; BAA13102.1; -.
CC DR EMBL: AB008796; AAL21070.1; -.
CC DR StyGene: SG10604; bglx.
CC DR InterPro: IPR002772; Glyco_hydro_3C.
CC DR InterPro: IPR001764; Glyco_hydro_3N.
CC Pfam: PF00933; Glyco_hydro_3; 1.
CC Pfam: PF01915; Glyco_hydro_3_C; 1.
CC PRINTS: PR00133; GHYDRLASE3.
CC DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
CC KW Hydrolase; Glycosidase; Peptidase; Signal; Complete proteome.
CC FT SIGNAL 1 20
CC FT CHAIN 1 765
CC FT ACT_SITE 287 287
CC FT CONFLICT 630 630
CC FT CONFLICT 634 635
CC FT CONFLICT 656 656
CC FT CONFLICT 674 674
CC FT CONFLICT 717 717
CC SQ SEQUENCE 765 AA; 83392 MW; C7C67BE6CAC96786 CRC64;

Query Match 38.2%; Score 42; DB 1; Length 765;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 PKGTQOYTDOIERPL 19
DB 624 PKKYSRYFDEANGPL 639

RESULT 15
MURI_STAM STANDARD; PRT; 266 AA.
ID MURI_STAM
QY 099UV6;
RC 28-FEB-2003 (Rel. 41, Created),
RC 28-FEB-2003 (Rel. 41, Last sequence update)
DE Glutamate racemase (EC 5.1.1.3).
DE MURI OR SAV1151 OR SA0997 OR MW1033.
OS Staphylococcus aureus (strain MU50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
CC Bacteria; Firmicutes; Bacillales; Staphylococcus.
CC NCBI_TaxID=158878; 158879; 196520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MU50 / ATCC 700699, and N315;
RC MEDLINE=2111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanemori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-U Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hizekawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanahita M., Yameshita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RT Lancet 357:1225-1240(2001).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RC MEDLINE=22040717; PubMed=12044378;
RC Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RT Lancet 359:1819-1827(2002).
CC -1- FUNCTION: Provides the (R)-glutamic acid required for cell wall
CC biosynthesis (By similarity).
CC -1- CATALYTIC ACTIVITY: L-glutamate = D-glutamate.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SIMILARITY: Belongs to the aspartate/glutamate racemases family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ibt-sib.ch/announce/
CC or send an email to license@ibt-sib.ch).
CC
CC EMBL: AP003361; BAB57313.1; -.
CC DR EMBL: AP003132; BAB42246.1; -.
CC DR EMBL: AP004825; BAB94898.1; -.
CC DR PIR: B89886; B89886.
CC DR HSSP: P56868; 1B74.
CC DR HAMAP: MF_00258; -.
CC DR InterPro: IPR001920; Asp/Glu_race.
CC DR InterPro: IPR004391; Glu_race.
CC Pfam: PF01177; Asp_Glu_race; 1.
CC TIGRPFAMs: TIGR00067; glut_race; 1.
CC DR PROSITE: PS00923; ASP_GLU_RACEMASE_1; 1.
CC DR PROSITE: PS00924; ASP_GLU_RACEMASE_2; 1.
CC DR PROSITE: PS00924; ASP_GLU_RACEMASE_2; 1.
CC KW Peptidoglycan synthetase; Cell wall; Isomerase; Complete proteome.
CC SQ SEQUENCE 266 AA; 29698 MW; 93BC35DBA42C7B0E CRC64;

Query Match 37.3%; Score 41; DB 1; Length 266;
Best Local Similarity 52.4%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 YFPNKGQ--QYTDIERPLM 20
DB 41 YGPRGEGVQYTVIRAKLM 61

RESULT 16
SPEE_BUCAL STANDARD; PRT; 286 AA.
ID SPEE_BUCAL
QY P57305;
RC 16-OCT-2001 (Rel. 40, Created)
RC 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Spermidine synthase (EC 2.5.1.16) (Putrescine aminopropyltransferase)
DE (SPSY).
GN SPEE OR BU209.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Buchnera.
CC NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RC MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS."
RT Nature 407:81-86(2000).
CC -1- CATALYTIC ACTIVITY: S-adenosylmethioninamine + putrescine = 5'-
CC methylthioadenosine + spermidine.
CC -1- PATHWAY: Spermidine biosynthesis; fifth (last) step.

```

```

CC -1- SIMILARITY: BELONGS TO THE SPERMIDINE/SPERMINE SYNTHASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB001118; BAB2926.1; -.
CC HAMAP; MF_00198; -; 1.
CC InterPro; IPR000051; SAM_bind.
CC InterPro; IPR001045; Spermine synthase.
CC Pfam; PF01564; Spermine synth; 1.
CC TIGRPFAM; TIGR00417; spse; 1.
CC PROSITE; PS01330; SPERMIDINE SYNTHASE; 1.
CC Spermidine biosynthesis; Transferrase; Complete proteome.
CC BINDING TO DECARBOXYLATED SAM
CC DOMAIN
CC 79 120
CC (POTENTIAL).
CC SEQUENCE 286 AA; 33788 MW; 5EE47D2107EB689E CRC64;

Query Match 37.3%; Score 41; DB 1; Length 286;
Best Local Similarity 58.3%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 YPNKGTQOQYTD 13
Db 120 YPNHNSQAYQD 131

RESULT 17
PYRD METJA STANDARD; PRT; 306 AA.
AC 058070;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dihydroorotate dehydrogenase (EC 1.3.3.1) (Dihydroorotate oxidase)
DE (DHODHase) (DHODase) (DHOD).
GN PYRD OR MJ0654.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=12190;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
MEDLINE=96337999; PubMed=8688087;
Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake O., Olsen G.J., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
Uiterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Frazer C.M., Smith H.O., Moese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RT Science 273:1058-1073(1996).
RL -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + O(2) = orotate +
H(2)O(2).
CC -1- COFACTOR: FMN (BY SIMILARITY).
CC -1- PATHWAY: Pyrimidine biosynthesis; fourth step.
CC -1- SUBUNIT: Heterotetramer of 2 pyrd and 2 pyrd subunits (By
similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the dihydroorotate dehydrogenase family.
Subfamily 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67513; AAB98649.1; -.
CC PIR; F64381; F64381.
CC HSSP; P54321; 2DOR.
CC TIGR; MJ0654; -.
CC HAMAP; MF_00224; -; 1.
CC InterPro; IPR001295; DHO_dh.
CC InterPro; IPR005720; DHO_dhl.
CC InterPro; IPR003009; FMN_enzyme.
CC Pfam; PF01180; DHODHase; 1.
CC TIGRPFAM; TIGR01037; pyrd sub1 fam; 1.
CC PROSITE; PS00912; DHODHASE_2; 1.
CC PROSITE; PS00911; DHODHASE_1; 1.
CC Pyrimidine biosynthesis; Oxidoreductase; Flavoprotein; FMN;
CC Complete proteome.
CC NP BIND 240 248
CC FMN (POTENTIAL).
CC SEQUENCE 306 AA; 32940 MW; B3532389BA9FC7C3 CRC64;

Query Match 37.3%; Score 41; DB 1; Length 306;
Best Local Similarity 42.9%; Pred. No. 25;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 4 PPKGTQYTDQIER 17
Db 76 PNPQVDEIERLEIK 89

RESULT 18
DDL LACLA STANDARD; PRT; 349 AA.
ID DDL LACLA
AC 09C15;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE D-alanine--D-alanine ligase (EC 6.3.2.4) (D-alanylanine synthetase)
DE (D-Ala-D-Ala ligase).
GN DDL OR LL0341.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=IL1403;
MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Winkler P., Manger S., Jallion O., Malarme K.,
Welschenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RL lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-
alanine-D-alanine.
CC -1- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; second
step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the D-alanine--D-alanine ligase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006271; AAK04439.1; -.
CC PIR; B86657; B86657.
CC HSSP; P07862; 11OW.
CC HAMAP; MF_00047; -; 1.

```


DR InterPro: IPR005905; D_alb_D_alb
DR InterPro: IPR000291; Dala_1lg Van.
DR Pfam: PF01820; Dala_D_liga; 1.
DR TIGFAMS: TIGR01205; D_alb_D_lattGR; 1.
DR PROSITE: PS00843; DALA_DALA_LIGASE_1; 1.
DR PROSITE: PS00844; DALA_DALA_LIGASE_2; FALSE NEG.
KW Ligase; Cell wall; Peptidoglycan synthesis; Complete proteome.
SQ SEQUENCE 349 AA; 38693 MW; 88E6996E1D2A8197 CRC64;

Query Match 36.8%; Score 40.5; DB 1; Length 349;
Best Local Similarity 47.4%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 2; Indels 5; Gaps 1.

OY 1 TYFPNKG-----TQYTTQ 14
||| ||| ||| ||| :
Db 40 TYFTTKGGEFTKTOEFTDK 58

RESULT 19
ATPG_SCHPO STANDARD; PRT; 301 AA.
● 074754; OSUSC2;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ATP synthase gamma chain, mitochondrial precursor (EC 3.6.3.14).
GN ATP3 OR SPBC1734.13.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
CX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RQ STRAIN=972;
RX MEDLINE=21848401; PubMed=11893360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Seguros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellon J., Simmonds R., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkhardt G., Aert R., Roben J., Grymoprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Filtz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Beger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "the genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE OF 27-125 FROM N.A.
RQ STRAIN=968 b90;
RX MEDLINE=20223868; PubMed=10755889;
RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA Hatake Y.;
RT "Large-scale screening of intracellular protein localization in living
fission yeast cells by the use of a GFP-fusion genomic DNA library.";
CC Genes Cells 5:169-190(2000).
-1- FUNCTION: Produces ATP from ADP in the presence of a proton
gradient across the membrane. The gamma chain is believed to be

```

CC important in regulating ATPase activity and the flow of protons
CC through the CF(0) complex. May play a role in the assembly/
CC stability of CF(1)'.
CC
CC -1 CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out)'.
CC
CC -1 SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: A, B and C (By similarity)'.
CC -1 SUBCELLULAR LOCATION: Mitochondrial (By similarity)'.
CC -1 SIMILARITY: Belongs to the ATPase gamma chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC DR EMBL, AL031856; CNA21307.1; -.
CC DR EMBL, AB027877; BAA67181.1; -.
CC DR PIR, T39660; T39660.
CC DR Genedb, SPombb: SPBC1734.13; -.
CC DR InterPro, IPR000131; ATPase_gamma.
CC DR Pfam, PF00231; ATP-synt. 1.
CC DR PRINTS, PR00126; ATPASEGAMMA.
CC DR TIGRFAMs, TIGR01146; ATPsyn_Fgamma. 1.
CC DR PROSITE, PS00153; ATPASE_GAMMA; 1.
CC KW ATP synthetase; CF(1); Hydrogen ion transport; Hydrolase;
CC Mitochondrion; Transic peptide.
CC FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL) .
CC FT CHAIN ? 301 ATP SYNTHASE GAMMA CHAIN.
CC ST SEQUENCE 301 AA; 33577 MW; 8595039B8E2F2B32 CRC64;
CC
CC
CC Query Match 36.4%; Score 40; DB 1; Length 301;
CC Best Local Similarity 35.3%; Pred. No. 36;
CC Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
CC
CC
CC OY 4 PNKGTOGYVDQIDIERPLM 20
CC ||| :||: |||
CC Db 217 PNLSAEYSDSEVHQPLM 233
CC
CC
CC RESULT 20
CC CAN1_RABIT
CC ID CAN1_RABIT STANDARD; PRT; 302 AA.
CC AC P06815;
CC DT 01-JAN-1988 (Rel. 06, Created)
CC DT 01-JAN-1988 (Rel. 06, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Calpain 1, large [catalytic] subunit (EC 3.4.22.17) [Calcium-activated
CC DE neutral proteinase] (CaNP) (Mu-type) (muCaNP) (Micromolar-calpain)
CC DE (Fragment).
CC DE CAPN1.
CC OS Oryctolagus cuniculus (Rabbit).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CC OC NCBI_TaxID=9986;
CC
CC RA NCB1
CC RA [1]
CC RA SEQUENCE FROM N.A.
CC RP MEDLINE=86250902; PubMed=2424911;
CC RA Emori Y., Kawasaki H., Sugihara H., Imajoh S., Kawashima S.,
CC RA Suzuki K.;
CC RT "Isolation and sequence analyses of cDNA clones for the large
CC RT subunits of two isozymes of rabbit calcium-dependent protease.";
CC RL J. Biol. Chem. 261:9465-9471(1986).
CC [2]
CC
CC RP CALCIUM-BINDING DATA.
CC RP MEDLINE=87279982; PubMed=3038855;
CC RA Minami Y., Emori Y., Kawasaki H., Suzuki K.;
CC RA "B-F" hand structure-domain of calcium-activated neutral protease
CC RA (CaNP) can bind Ca2+ ions.";

```

```

RL 5. Biochem. 101:889-895(1987).
CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
CC catalyze limited proteolysis of substrates involved in
CC cytoskeletal remodelling and signal transduction.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
CC Arg-Xaa with Leu or Val as the P2 residue.
CC -1- COFACTOR: Binds 3 calcium ions (By similarity).
CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
CC calcium and inhibited by calpastatin.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC membrane upon Ca++ binding (By similarity).
CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M3363; AAA31456.1; -
CC PIR; A24815; A24815.
CC HSSP; 007009; IDPO.
CC MEROPS; C02.001; -
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR001300; Protease_C2.
CC InterPro; IPR000169; SHprot_acetate.
CC Pfam; PF01067; Calpain_III; 1.
CC SMART; SM00036; efhand_3.
CC SMART; SM00720; calpain_III; 1.
CC SMART; SM00054; Efh; 3.
CC PROSITE; PS00018; EF_HAND; 2.
CC PROSITE; PS00139; THIOL_PROTEASE_CYS; PARTIAL.
CC PROSITE; PS00639; THIOL_PROTEASE_HIS; PARTIAL.
CC PROSITE; PS00640; THIOL_PROTEASE ASN; PARTIAL.
CC Hydroxylase; Thiol protease; Calcium-binding; Multigene family.
CC NON TER
CC DOMAIN 1 1
CC FT DOMAIN <1 114 DOMAIN III.
CC FT DOMAIN 115 130 LINKER.
CC FT DOMAIN 131 301 DOMAIN IV.
CC FT CA BIND 142 153 EF-HAND 1.
CC FT CA BIND 186 197 EF-HAND 2.
CC FT CA BIND 216 227 EF-HAND 3.
CC DOMAIN 251 262 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
CC DOMAIN 281 292 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
CC SEQUENCE 302 AA; 35274 MW; 178BPEF4216CEAB CRC64;

Query Match 36.4%; Score 40; DB 1; Length 302;
Best Local Similarity 33.3%; Pred. No. 36;
Matches 11; Conservative 3; Mismatches 5; Indels 14; Gaps 1;

QY 1 TYFPNK-----GTQGYDQIERPL 19
DB 88 TFEPNKGGDFVLFKFEKKGAGTGLDQGANL 120

RESULT 21
MOAA_BACHD STANDARD; PRT; 338 AA.
AC 09K9M9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Molybdenum cofactor biosynthesis protein A.
GN MOAA OR NARA OR BH2526.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=86665;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhlara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: Together with moac, is involved in the conversion of a
CC guanosine derivative (GMP) into molybdopterin precursor Z (By
CC similarity).
CC -1- COFACTOR: Binds 1 3Fe-4S cluster (By similarity).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis; first step.
CC -1- SIMILARITY: BELONGS TO THE MOAA / NIFB / POOB FAMILY. MOAA
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001515; BAB06245.1; -
CC PIR; P83965; P83965.
CC HAMAP; MF_01225; -; 1.
CC InterPro; IPR006638; ELP3.
CC InterPro; IPR000385; MOAA_NiFe_Pqge.
CC Pfam; PF04055; Radical_SAM; 1.
CC SMART; SM00729; ELP3; 1.
CC PROSITE; PS01305; MOAA_NiFe_POOB; 1.
CC Molybdenum cofactor biosynthesis; Metal-binding; Iron; Iron-sulfur;
CC 3Fe-4S; Complete proteome.
CC KW METAL 25 25 IRON-SULFUR (3FE-4S) (POTENTIAL).
CC FT METAL 32 32 IRON-SULFUR (3FE-4S) (POTENTIAL).
CC FT METAL 263 263 IRON-SULFUR (3FE-4S) (POTENTIAL).
CC FT METAL 280 280 IRON-SULFUR (3FE-4S) (POTENTIAL).
CC SO SEQUENCE 338 AA; 38157 MW; 910258583C411B95 CRC64;

Query Match 36.4%; Score 40; DB 1; Length 338;
Best Local Similarity 58.3%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 GTQGYDQIERP 18
DB 2 GTQPYDRLNRP 13

RESULT 22
NI4M_CHICK STANDARD; PRT; 459 AA.
AC P18939;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN MTND4 OR ND4 OR NADH4.
OS Gallus gallus (Chicken).
OS Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_Taxid=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90230301; PubMed=2329578;
RA Desjardins P., Morais R.;
RT "Sequence and gene organization of the chicken mitochondrial genome.
RT A novel gene order in higher vertebrates.";
RL J. Mol. Biol. 212:599-634(1990).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

```

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X52392; CAA36634.1; -
DR PIR; S10196; S10196.
DR InterPro; IPR003918; NADHox_oxred4.
DR InterPro; IPR001750; Oxidored_g1.
DR InterPro; IPR000260; Oxidored_g5_N.
DR Pfam; PF00361; oxidored_g1; 1.
DR Pfam; PF01059; oxidored_g5_N; 1.
DR PRINTS; PR01437; NUOXDRPTA54.
KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 459 AA; 51145 MW; 46D140070B9D648A CRC64;

Query Match 36.4%; Score 40; DB 1; Length 459;
Best Local Similarity 40.9%; Pred. No. 57;
Matches 9; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 1 TYPENKGTQQT-DQIERPLM 20
Db 45 SYPTKTLTWTGMDISTPL 66

RESULT 23
NAP MOUSE STANDARD; PRT; 563 AA.
ID NAP_MOUSE
AC Q9CWL8; Q8VHE3;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Nuclear associated protein (NAP).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RX MEDLINE=10090;
RA NIKAIKO I., Ohtsuka N., Saito R., Suzuki H., Yamana K., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
RA Balderelli R., Hall D.P., Bull C., Hume D.A., Quackenbush J.,
RA Schiml L.M., Kanapin A., Matveeva H., Batalov S., Beisel K.W.,
RA Blake J.A., Brad D., Bruscia V., Chochia C., Corbani L.B., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King P.A.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenle L., Miki H.,
RA Nagashima T., Numa K., Okido T., Pavan W.J., Pereira G., Pesole G.,
RA Petrovsky N., Piliat R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi F., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wysshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carrinci P., Hayatsu N.,
RA Hitzane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Aizawa T., Fukuda S.,

```

```

RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
CC - SUBCELLULAR LOCATION: Nuclear (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AY009405; AAK7389.1; -
DR EMBL; AK010547; BAB27020.1; -
DR EMBL; AK044690; BAC32034.1; -
DR EMBL; AK077616; BAC36902.1; -
DR MGD; MGI:1913892; 5730471K09R1K.
KM Nuclear protein.
FT DOMAIN 22 25 POLY-GLU.
FT DOMAIN 68 75 POLY-GLU.
FT CONFLICT 121 121 I -> V (IN REF. 2; BAC32034).
FT CONFLICT 399 400 TE -> OO (IN REF. 1).
FT CONFLICT 418 418 L -> M (IN REF. 2; BAC36902).
SQ SEQUENCE 563 AA; 64979 MW; 89B67CF3A379C3D2 CRC64;

Query Match 36.4%; Score 40; DB 1; Length 563;
Best Local Similarity 42.1%; Pred. No. 71;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 TYPENKGTQQT-DQIERPL 19
Db 8 SYPTKTLTWTGMDISTPL 26

RESULT 24
NAP MOUSE STANDARD; PRT; 713 AA.
ID NAP_MOUSE
AC Q35350; O88666;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Calpain 1, large [catalytic] subunit (EC 3.4.22.17) (Calcium-activated
DE neutral proteinase) (CAPN) (Mu-type) (muCAPN) (Microtubule-calpain).
GN CAPN1 OR CAPN1 OR CAPN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RX MEDLINE=10090;
RA STRAIN=CP-1;
RA Pottier C., Pousard S., Faust D.M., Imaizumi-Sherer T., Weiss M.C.,
RA Ducasting A., Montarras D., Pinet C., Guenet J.-P.,
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
(2)
RA SEQUENCE FROM N.A.
RA STRAIN=129;
RA Sahr K.E., Andrabi S., Peters L.L., Chiefti A.H.;
RL "Cloning and characterization of the cDNA and gene encoding the mouse
RL mu-calpain large subunit protein."
CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
CC catalyzes limited proteolysis of substrates involved in
CC cytoskeletal remodeling and signal transduction.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-[Xaa, Met]-[Xaa or
CC Arg]-[Xaa with Leu or Val as the P2 residue.
CC -1- COFACTOR: Binds 3 calcium ions (by similarity).
CC -1- ENZYME REGULATION: Activated by micromolar concentrations of

```

```

CC      calcium and inhibited by calpastatin.
CC      -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC      (regulatory) subunit.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC      membrane upon Ca++ binding (By similarity).
CC      -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF021847; AAB72222.1; -.
DR      EMBL; AF084459; AAC3134.1; -.
DR      HSSP; Q07009; IDPO.
DR      MGDP; MG188263; Capn1.
DR      InterPro; IPR002048; EF-hand.
DR      InterPro; IPR001300; Protease_C2.
DR      InterPro; IPR000169; SHprot_acsite.
DR      Pfam; PF01067; Calpain_III; 1.
DR      Pfam; PF00036; ehand; 3.
DR      Pfam; PF00648; Peptidase_C2; 1.
DR      PRINTS; PR00704; CALPAIN_III; 1.
DR      SMART; SM00720; calpain_III; 1.
DR      SMART; SM00230; Cyspc; 1.
DR      SMART; SM00054; Eph; 3.
DR      PROSITE; PS00018; EF_HAND; 2.
DR      PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR      PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
DR      PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
DR      HydroLase; Thiol protease; Calcium-binding; Multigene family.
KW      DOMAIN
FT      DOMAIN 30 365
FT      DOMAIN 366 525
FT      DOMAIN 526 541
FT      DOMAIN 542 541
FT      DOMAIN 542 712
FT      CA_BIND 553 564
FT      CA_BIND 553 564
FT      CA_BIND 597 608
FT      CA_BIND 627 638
FT      CA_BIND 627 638
FT      DOMAIN 662 673
FT      DOMAIN 692 703
FT      ACT_SITE 115 115
FT      ACT_SITE 272 272
FT      ACT_SITE 272 272
FT      ACT_SITE 296 296
FT      ACT_SITE 488 488
FT      ACT_SITE 696 696
FT      ACT_SITE 713 AA; 82106 MW; 3E1E26C95802B864 CRC64;
SQ      SEQUENCE
SQ      Query Match 36.4%; Score 40; DB 1; Length 713;
SQ      Best Local Similarity 33.3%; Pred. No. 93;
SQ      Matches 11; Conservative 3; Mismatches 5; Indels 14; Gaps 1;
OY      1 TYFPNK-----GTQOYTDQIERPL 19
OY      499 TFEFNKEDPLRPFSEKKAQTGLDDQIOANL 531
DB

```

```

OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxId=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=97107433; PubMed=8950173;
RA      Sotomachi H., Ameno S., Ishiura S., Suzuki K.;
RT      "Primary sequences of rat mu-calpain large and small subunits are,
RT      respectively, moderately and highly similar to those of human.";
RL      Biochim. Biophys. Acta 1309:37-41(1996).
CC      -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
CC      catalyze limited proteolysis of substrates involved in
CC      cytoskeletal remodeling and signal transduction.
CC      -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Met-|-Xaa or
CC      Arg-|-Xaa with Leu or Val as the P2 residue.
CC      -1- COFACTOR: Binds 3 calcium ions (By similarity).
CC      -1- ENZYME REGULATION: Activated by micromolar concentrations of
CC      calcium and inhibited by calpastatin.
CC      -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC      (regulatory) subunit.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC      membrane upon Ca++ binding (By similarity).
CC      -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U53858; AAC53001.1; -.
DR      PDB; 1KKR; 20-MAR-02.
DR      MGDP; C02.001; -.
DR      InterPro; IPR002048; EF-hand.
DR      InterPro; IPR001300; Protease_C2.
DR      InterPro; IPR000169; SHprot_acsite.
DR      Pfam; PF01067; Calpain_III; 1.
DR      Pfam; PF00036; ehand; 3.
DR      Pfam; PF00648; Peptidase_C2; 1.
DR      PRINTS; PR00704; CALPAIN.
DR      SMART; SM00720; calpain_III; 1.
DR      SMART; SM00230; Cyspc; 1.
DR      SMART; SM00054; Eph; 3.
DR      PROSITE; PS00018; EF_HAND; 2.
DR      PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR      PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
DR      PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
DR      Calcium-binding; HydroLase; Thiol protease; 3D-structure.
KW      DOMAIN
FT      DOMAIN 30 365
FT      DOMAIN 366 525
FT      DOMAIN 526 541
FT      DOMAIN 542 541
FT      DOMAIN 542 712
FT      CA_BIND 553 564
FT      CA_BIND 597 608
FT      CA_BIND 627 638
FT      CA_BIND 627 638
FT      DOMAIN 662 673
FT      DOMAIN 692 703
FT      ACT_SITE 115 115
FT      ACT_SITE 272 272
FT      ACT_SITE 296 296
FT      ACT_SITE 488 488
FT      ACT_SITE 696 696
FT      ACT_SITE 713 AA; 82118 MW; 6E6646080BEFAB6 CRC64;
SQ      SEQUENCE
SQ      Query Match 36.4%; Score 40; DB 1; Length 713;
SQ      Best Local Similarity 33.3%; Pred. No. 93;
SQ      Matches 11; Conservative 3; Mismatches 5; Indels 14; Gaps 1;
OY      1 TYFPNK-----GTQOYTDQIERPL 19
OY      499 TFEFNKEDPLRPFSEKKAQTGLDDQIOANL 531
DB

```

RESULT 26
GUNG_CLOCE STANDARD; PRT; 725 AA.
AC P37700;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Endoglucanase G precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase G)
GN (Cellulase G) (EGCGG).
OS Clostridium cellulolyticum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1521;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 35319;
RX MEDLINE=93012971; PubMed=1398087;
RA Bagmar-Tardif C., Gaudin C., Belaich A., Hoest P., Citard T., Belaich J.-P.;
"Sequence analysis of a gene cluster encoding cellulases from Clostridium cellulolyticum.";
Gene 119:17-28(1992).
RL
CC -1- FUNCTION: THE BIOLOGICAL CONVERSION OF CELLULOSE TO GLUCOSE
CC GENERALLY REQUIRES THREE TYPES OF HYDROLYTIC ENZYMES:
CC (1) ENDOLUCANASES WHICH CUT INTERNAL BETA-1,4-GLUCOSIDIC BONDS;
CC (2) EXOCELLULOBIODHROLASES THAT CUT THE DISACCHARIDE CELLULOSE
CC FROM THE NONREDUCING END OF THE CELLULOSE POLYMER CHAIN;
CC (3) BETA-1,4-GLUCOSIDASES WHICH HYDROLYSE THE CELLULOSE AND OTHER
CC SHORT CELLO-OLIGOSACCHARIDES TO GLUCOSE.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC linkages in cellulose, lichenin and cereal beta-D-glucans.
CC -1- PATHWAY: Cellulose degradation.
CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
CC WELL AS IN OTHER C.THERMOCELLUM CELLULOSE ENZYMES. THIS DOMAIN
CC MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY E (FAMILY 9 OF GLYCOSYL
CC HYDROLASES).
CC -1- SIMILARITY: Contains 1 type-3 cellulose-binding (CBD) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M87018; AAA73868.1; -;
CC PIR; JCI300; JCI300.
CC HSSP; P26221; 1TF4.
DR InterPro; IPR001956; CBD 3.
DR InterPro; IPR002105; Dockerin_1.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001701; Glyco_hydro_9.
DR Pfam; PF00942; CBM 3; 1.
DR Pfam; PF00404; Dockerin_1; 2.
DR Pfam; PF00759; Glyco_hydro_9; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
DR PROSITE; PS00448; CLOS_CELULOSOME_RPT; 1.
DR PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
DR PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
KW Cellulose degradation; Hydrolase; Glycosidase; Repeat; Signal.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 725 ENDOLUCANASE G.
FT DOMAIN 485 662 CELLULOSE-BINDING (BY SIMILARITY).
FT ACT SITE 408 408 BY SIMILARITY.
FT ACT SITE 446 446 BY SIMILARITY.
FT ACT SITE 455 455 BY SIMILARITY.
FT DOMAIN 664 719 2 X 24 AA APPROXIMATE REPEATS.
FT REPEAT 664 687 1.
FT REPEAT 696 719 2.
FT SEQUENCE 725 AA; 79886 MW; EC81F84F93A6872C CRC64;

Query Match 36.4%; Score 40; DB 1; Length 725;
Best Local Similarity 53.3%; Pred. No. 94;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Cy 1 TYPENKGTQGYTDQI 15
Db 271 SYFPMWKEQQTDLI 285
RESULT 27
LOXX_SOYBN STANDARD; PRT; 864 AA.
ID LOXX_SOYBN
AC P24095;
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Seed lipoxigenase (EC 1.13.11.12).
GN LOX1.4 OR SC514.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eufrosida I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RX MEDLINE=91370880; PubMed=1909908;
RA Shibata D., Kato T., Tanaka K.;
RT "Nucleotide sequences of a soybean lipoxigenase gene and the short
RT intergenic region between an upstream lipoxigenase gene.";
RL Plant Mol. Biol. 16:353-359(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams 82; TISSUE=Radicle;
RA Park T., Holland M.A., Laeskey J.G., Polacco J.C.;
RL Submitted (DEC-1993) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PLANT LIPOXYGENASE MAY BE INVOLVED IN A NUMBER OF
CC DIVERSE ASPECTS OF PLANT PHYSIOLOGY INCLUDING GROWTH AND
CC DEVELOPMENT, PEST RESISTANCE, AND SENESENCE OR RESPONSES TO
CC WOUNDING. IT CATALYZES THE HYDROPEROXIDATION OF LIPIDS, CONTAINING
CC A CIS, CIS-1,4-PENTADIENE STRUCTURE.
CC -1- CATALYTIC ACTIVITY: Linoleate + O(2) = (9Z,11E)-(13S)-13-
CC hydroperoxyoctadeca-9,11-dienoate.
CC -1- COFACTOR: IRON, ONE ATOM TIGHTLY BOUND PER MOLECULE.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: GERMINATED COTYLEDONS.
CC -1- INDUCTION: By jasmonate.
CC -1- MISCELLANEOUS: SOYBEAN CONTAINS AT LEAST 4 DISTINCT ISOENZYMES,
CC L-1, L-2, L-3A AND L-3B IN DRY SEEDS, AND AT LEAST TWO DISTINCT
CC ISOZYMES IN THE HYPOCOTYL/RADICLE REGION OF THE SEEDLING STEM.
CC -1- SIMILARITY: BELONGS TO THE LIPOXYGENASE FAMILY.
CC -1- SIMILARITY: Contains 1 PLAT domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56139; CAJ39604.1; -;
CC EMBL; U04526; AAA03728.1; -;
CC PIR; S13381; S13381.
CC HSSP; P08170; 2SBL.
DR InterPro; IPR000907; Lipoxigenase.
DR InterPro; IPR001024; Lipoxigenase_LH2.
DR Pfam; PF00305; Lipoxigenase; 1.
DR Pfam; PF01477; PLAT; 1.
DR PRINTS; PR00087; LIPOXYGENASE.
DR SMART; SM00308; LH2; 1.

```

DR PROSITE; PS00711; LIPOXYGENASE_1; 1.
DR PROSITE; PS00081; LIPOXYGENASE_2; 1.
DR PROSITE; PS00095; PLAT; 1.
KM Oxidoreductase; Dioxigenase; Iron; Multigene family.
FT DOMAIN 44 171
FT METAL 525 525 IRON (BY SIMILARITY).
FT METAL 530 530 IRON (BY SIMILARITY).
FT METAL 716 716 IRON (BY SIMILARITY).
FT METAL 864 864 IRON (BY SIMILARITY).
FT CONFLICT 233 233 S -> C (IN REF. 2).
FT CONFLICT 240 240 R -> L (IN REF. 2).
FT CONFLICT 604 604 D -> H (IN REF. 2).
FT CONFLICT 695 695 M -> K (IN REF. 2).
SQ SEQUENCE 864 AA; 96816 MW; 66F31FBI5AF5B60 CRC64;

Query Match 36.4%; Score 40; DB 1; Length 864;
Best Local Similarity 29.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

2 YFPNKGTOOYTDQIERP 18
: ||| : : :
740 FIPAEETPEYDEMKNP 756

RESULT 28
HYSA_STRPN STANDARD; PRT; 1066 AA.
AC 054873; 054874; 37, Created)
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-DEC-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hyaluronate lyase precursor (EC 4.2.2.1) (Hyaluronidase) (Hyase).
GN SP0314.
OC Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxId=1313;
[1] _SEQUENCE FROM N.A.
RP STRAIN=ATCC BAA-334 / TIGR4;
RC MEDLINE=21357209; PubMed=1163916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Bisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radue D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RA "Complete genome sequence of a virulent isolate of Streptococcus
RA pneumoniae.";
RA Science 293:498-506(2001).
RL [2]
RN SEQUENCE OF 89-1066 FROM N.A.
RP STRAIN=Type 23;
RC MEDLINE=94156460; PubMed=8112843;
RA Berry A.M., Lock R.A., Thomas S.M., Rajan D.P., Hansman D.,
RA Paton J.C.;
RA "Cloning and nucleotide sequence of the Streptococcus pneumoniae
RA hyaluronidase gene and purification of the enzyme from recombinant
RA Escherichia coli.";
RA Infect. Immun. 62:1101-1108(1994).
RN [3]
RP CRYSTALLIZATION.
RP MEDLINE=98234706; PubMed=9573623;
RA Jedrejdas M.J., Chantelat L., Mewbourne R.B.;
RT "Crystallization and preliminary X-ray analysis of Streptococcus
RT pneumoniae hyaluronate lyase.";
RT J. Struct. Biol. 121:73-75(1998).
CC -1- CATALYTIC ACTIVITY: Hyaluronate = N 3-(4-deoxy-beta-D-gluc-4-
CC enuronosyl)-N-acetyl-D-glucosamine.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan
CC by an amide bond (Potential).

```

```

CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB007344; AAK74491.1; -.
DR EMBL; L20670; AAK53685.1; ALT INIT.
DR EMBL; L20670; AAK53686.1; ALT INIT.
DR PIR; B95037; B95037.
DR PDB; 1C82; 05-APR-01.
DR PDB; 1P9G; 16-MAY-01.
DR PDB; 1LOH; 07-AUG-02.
DR PDB; 1LXK; 07-AUG-02.
DR PDB; 1N7N; 31-DEC-02.
DR PDB; 1N7O; 31-DEC-02.
DR PDB; 1N7P; 31-DEC-02.
DR PDB; 1N7Q; 31-DEC-02.
DR PDB; 1N7R; 31-DEC-02.
DR TIGR; SP0314; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR006192; LPXTG.
DR InterPro; IPR003159; Lyase_8.
DR InterPro; IPR004103; Lyase_8_C.
DR Pfam; PF02018; CBM_4_9; 1.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02278; Lyase_8; 1.
DR Pfam; PF02884; Lyase_8_C; 1.
DR TIGRAME; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; FALSE NEG.
KW Lyase; Cell wall; Peptidoglycan-anchor; Signal; Complete proteome;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 1039
FT PROPEP 1040 1066
FT SITE 1036 1040
FT MOD_RES 1039 1039
FT CONFLICT 108 108 H -> P (IN REF. 2).
FT CONFLICT 115 115 L -> V (IN REF. 2).
FT CONFLICT 139 139 I -> V (IN REF. 2).
FT CONFLICT 211 211 L -> P (IN REF. 2).
FT CONFLICT 236 236 S -> P (IN REF. 2).
FT CONFLICT 290 290 A -> T (IN REF. 2).
FT CONFLICT 313 313 E -> D (IN REF. 2).
FT CONFLICT 340 340 T -> I (IN REF. 2).
FT CONFLICT 613 613 C -> R (IN REF. 2).
FT CONFLICT 658 658 P -> T (IN REF. 2).
FT CONFLICT 658 658 G -> S (IN REF. 2).
FT CONFLICT 821 821 V -> G (IN REF. 2).
FT CONFLICT 848 848 F -> S (IN REF. 2).
FT CONFLICT 853 853 R -> G (IN REF. 2).
FT CONFLICT 907 907 R -> G (IN REF. 2).
SQ SEQUENCE 1066 AA; 120771 MW; 81DB22A837BE61F9 CRC64;

Query Match 36.4%; Score 40; DB 1; Length 1066;
Best Local Similarity 43.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

2 YFPNKGTOOYTDQIERP 17
||| : : :
429 YFSDEIKYTDVIEK 444

RESULT 29
PER2_MOUSE STANDARD; PRT; 1257 AA.
AC 054943; 054954;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

DE Period circadian protein 2 (mper2).
 GN PER2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxId=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98088803; PubMed=9428527;
 RA Albrecht U., Sun Z.S., Rieche G., Lee C.C.;
 RT "A differential response of two putative mammalian circadian
 RL regulators, mper1 and mper2, to light.";
 RN Cell 91:1055-1064 (1997).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98087121; PubMed=9427249;
 RA Shearman L.P., Zylka M.J., Weaver D.R., Kolakowski L.F. Jr.,
 RA Reppert S.M.;
 RT "Two period homologs: circadian expression and photic regulation in
 the suprachiasmatic nuclei.";
 RN Neuron 19:1261-1269 (1997).
 [3]
 RP REVISIONS TO 172 AND 501.
 RA Zylka M.J., Reppert S.M.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CIRCADIAN REGULATOR THAT MAY ACT AS A TRANSCRIPTION
 CC FACTOR. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL
 CC LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT
 CC TRANSCRIPTIONAL INHIBITION. EXPRESSION OSCILLATES WITH A 24 HR
 CC RHYTHM IN THE SUPRACHIASMATIC NUCLEUS (SCN) AND THE WHOLE EYES.
 CC OSCILLATIONS ARE MAINTAINED UNDER CONSTANT DARKNESS AND ARE
 CC RESPONSIVE TO CHANGES OF THE LIGHT/DARK CYCLES. THERE IS A 4 HOUR
 CC TIME DELAY BETWEEN PER1 AND PER2 OSCILLATIONS. THE EXPRESSION
 CC RHYTHMS APPEAR TO ORIGINATE FROM RETINA.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- TISSUE SPECIFICITY: IN THE BRAIN, HIGH EXPRESSION IN SCN DURING
 CC THE SUBJECTIVE DAY. CONSTITUTIVE EXPRESSION IN THE CORNU AMMONIS
 CC AND IN THE DENTATE GYRUS OF THE HIPPOCAMPUS. ALSO EXPRESSED IN THE
 CC PERIFORM CORTEX AND THE GLOMERULI OF THE OLFACTORY BULB, AND AT A
 CC LOWER EXTENT IN THE CEREBRAL CORTEX. NOT EXPRESSED IN THE PARS
 CC TUBERALIS AND THE POKINJIE NEURONS. ALSO FOUND IN HEART, BRAIN,
 CC LUNG, LIVER, KIDNEY. AT LOW LEVEL IN SPLEEN, BUT NOT IN SKELETAL
 CC MUSCLE OR TESTIS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE SCN DURING LATE FETAL AND
 CC EARLY NEONATAL LIFE.
 CC -1- INDUCTION: BY LIGHT EXPOSURE DURING SUBJECTIVE NIGHT, BUT NOT
 CC DURING SUBJECTIVE DAY.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS.
 CC -1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF036893; AAC39942.1; -;
 CC EMBL; AF035830; AAC3592.1; -;
 CC PIR; T09493; T09493.
 CC MGI; MGI:1195265; Per2.
 CC InterPro; IPR001610; PAC.
 CC InterPro; IPR000014; PAS_domain.
 CC SMART; SMO0086; PAC; 1.
 CC SMART; SMO0091; PAS; 1.
 CC PROSITE; PS00112; PAS; 1.
 CC Transcription Regulation; Nuclear protein; Repeat; Biological rhythms.
 CC DOMAIN 107 144 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 CC DOMAIN 180 246 PAS 1.

FT DOMAIN 317 387 PAS 2.
 FT DOMAIN 396 436 PAC.
 FT DOMAIN 778 795 NUCLEAR LOCALIZATION SIGNAL.
 FT CONFLICT 445 445 P -> S (IN REF. 2).
 FT CONFLICT 728 728 K -> R (IN REF. 2).
 SQ SEQUENCE 1257 AA; 135881 MW; 5548BAFF036CF7B CRC64;
 Query Match 36.4%; Score 40; DB 1; Length 1257;
 Best Local Similarity 47.1%; Pred. No. 1.7e+02;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 4 PNKGQYTDQIERPLM 20
 DB 453 PHSVQELTEQIHRLLM 469
 RESULT 30
 SUI5_RABIT
 ID SUI5_RABIT STANDARD; PRT; 1826 AA.
 AC P07768;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Sucrase-isomaltase, intestinal [contains: Sucrase (EC 3.2.1.48);
 DE Isomaltase (EC 3.2.1.10)].
 GN SI.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NX NCBI_TaxId=9986;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86245069; PubMed=3755079;
 RA Hunziker W., Spiess M., Semenza G., Lodish H.F.;
 RT "The sucrase-isomaltase complex: primary structure, membrane-
 RT orientation, and evolution of a stalked, intrinsic brush border
 RL protein.";
 RN Cell 46:227-234 (1986).
 [2]
 RP PRELIMINARY SEQUENCE OF 1-37 AND 1007-1014.
 RX MEDLINE=83105704; PubMed=7152027;
 RA Sjoestrom H., Noren O., Christensen L.A., Wacker H., Spiess M.,
 RA Bigler-Meyer B., Rickli E.B., Semenza G.;
 RT "N-terminal sequences of pig intestinal sucrase-isomaltase and pro-
 RT sucrase-isomaltase. Implications for the biosynthesis and membrane
 RT insertion of pro-sucrase-isomaltase.";
 RL FEBS Lett. 148:321-325 (1982).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE FINAL STAGE OF
 CC CARBOHYDRATE DIGESTION.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of sucrose and maltose by an
 CC alpha-D-glucosidase-type action.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,6-alpha-D-glucosidic linkages
 CC in some oligosaccharides produced from starch and glycogen by
 CC alpha-amylase, and in isomaltose.
 CC -1- SUBUNIT: THE RESULTING SUCRASE AND ISOMALTASE SUBUNITS STAY
 CC ASSOCIATED WITH ONE ANOTHER IN A COMPLEX BY NON-COVALENT LINKAGES.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Brush border.
 CC -1- PTM: THE PRECURSOR IS PROTEOLYTICALLY CLEAVED WHEN EXPOSED TO
 CC PANCREATIC PROTEASES IN THE INTESTINAL LUMEN.
 CC -1- PTM: N- AND O-GLYCOSYLATED.
 CC -1- MISCELLANEOUS: THERE IS A HIGH DEGREE OF HOMOLOGY BETWEEN THE
 CC ISOMALTASE AND SUCRASE PORTIONS (41 % OF AMINO ACID IDENTITY)
 CC INDICATING THAT THIS PROTEIN IS EVOLVED BY PARTIAL GENE
 CC DUPLICATION.
 CC -1- SIMILARITY: Contains 1 P-type (cref11) domain.
 CC -1- SIMILARITY: BELONGS TO FAMILY 31 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL, M4046; AAA1459.1. -

CC InterPro; IPR000322; Glyco_hydro_31.

CC InterPro; IPR000519; P_trefoil.

CC Pfam; PF01055; Glyco_hydro_31; 2.

CC Pfam; PF00088; trefoil; 2.

CC SMART; SM00018; PD; 2.

CC PROSITE; PS00025; P_TREFOIL; 1.

CC PROSITE; PS00129; GLYCOSYL_HYDROL_F31_1; 2.

CC PROSITE; PS00707; GLYCOSYL_HYDROL_F31_2; 2.

CC Multifunctional enzyme; Transmembrane; Glycoprotein; Hydrolase;

CC Glycosidase; Repeat; Signal-anchor; Sulfation.

FT INIT MET 0

FT CHAIN 1 1826 SUCRASE-ISOMALTASE, INTERSTITIAL.

FT CHAIN 1 1006 ISOMALTASE.

FT CHAIN 1007 1826 SUCRASE.

FT TRANSMEM 1 11 CYTOPLASMIC.

FT DOMAIN 12 31 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).

FT DOMAIN 42 59 LIMENAL.

FT DOMAIN 61 108 SER/TR-RICH.

FT DOMAIN 109 1006 P-TYPE.

FT DOMAIN 1007 1826 ISOMALTASE.

FT ACT_SITE 504 504 SUCRASE.

FT ACT_SITE 1393 1393

FT DISULFID 62 93

FT DISULFID 76 92

FT DISULFID 87 105 BY SIMILARITY.

FT MOD_RES 350 390 Sulfation (POTENTIAL).

FT MOD_RES 1381 1381 Sulfation (POTENTIAL).

FT MOD_RES 1384 1384 Sulfation (POTENTIAL).

FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 858 858 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 895 895 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 903 903 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1001 1001 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1234 1234 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1302 1302 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1324 1324 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1339 1339 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1353 1353 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1367 1367 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1402 1402 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1534 1534 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1571 1571 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1747 1747 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1762 1762 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1798 1798 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT SEQUENCE 1826 AA; 210008 MW; 6840D03955A45B55 CRC64;

Query Match 36.4%; Score 40; DB 1; Length 1826;

Best Local Similarity 47.1%; Pred. No. 2.6e+02;

Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 4 PNRGQYTDQIERPLM 20

Db 1201 PEVATQYHEVIGHPRM 1217

RESULT 31

T2D1 HUMAN STANDARD; PRT; 1872 AA.

AC P21675;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Transcription initiation factor TFIID 250 kDa subunit (TAFII-250)

DE (TAFII250) (TBP-associated factor 250 kDa) (P250) (Cell cycle gene 1

protein).

GN TAFI OR TAF2A OR CCGI OR BA2R.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_Taxid=9606;

XP [1]

RC TISSUE=lymphoid carcinoma;

RX MEDLINE=91246200; PubMed=2038334;

RA Sekiguchi T., Nozito Y., Nakamura Y., Hisamoto N., Nishimoto T.;

RT "The human CCGI gene, essential for progression of the G1 phase,

RT encodes a 210-kilodalton nuclear DNA-binding protein.";

RL Mol. Cell. Biol. 11:3317-3325(1991).

RN [2]

RP PRELIMINARY SEQUENCE FROM N.A.

RX MEDLINE=8905056; PubMed=3169001;

RA Sekiguchi T., Miyata T., Nishimoto T.;

RT "Molecular cloning of the cDNA of human X chromosomal gene (CCGI)

RT which complements the temperature-sensitive G1 mutants, tsBN462 and

RL ts1, of the BHK cell line.";

RL EMO J. 7:1683-1687(1988).

RN [3]

RP SEQUENCE FROM N.A. AND CHARACTERIZATION.

RX MEDLINE=93196704; PubMed=7680771;

RA Ruppert S., Wang B.H., Tjian R.;

RT "Cloning and expression of human TAFII250: a TBP-associated factor

RT implicated in cell-cycle regulation.";

RL Nature 362:115-119(1993).

RN [4]

RP CHARACTERIZATION.

RX MEDLINE=93196705; PubMed=8450888;

RA Hiestake K., Hasegawa S., Takada R., Nakatani Y., Horikoshi M.,

RA Roeder R.G.;

RT "The p250 subunit of native TATA box-binding factor TFIID is the

RT cell-cycle regulatory protein CCGI.";

RL Nature 362:179-181(1993).

CC -1- FUNCTION: MAY PLAY AN ESSENTIAL ROLE IN TFIID ASSEMBLY BY

CC INTERACTING WITH BOTH TBP AND OTHER TAF. AS WELL AS SERVING TO

CC LINK THE CONTROL OF TRANSCRIPTION TO THE CELL CYCLE. ESSENTIAL

CC FOR PROGRESSION OF THE G1 PHASE OF THE CELL CYCLE. POSSESSES

CC DNA-BINDING ACTIVITY.

CC -1- SUBUNIT: TFIID IS COMPOSED OF TBP AND A VARIETY OF TBP-ASSOCIATED

CC FACTORS.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- PTM: PHOSPHORYLATED BY CASEIN KINASE II IN VITRO.

CC -1- SIMILARITY: Contains 2 bromodomains.

CC -1- SIMILARITY: Contains 1 HMG box domain.

CC -1- SIMILARITY: TO DROSOPHILA TAFII-230. SOME TO S.POMBE TAFII-111

CC AND TO S.CEREVISIAE TAF145.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL, D90359; BAA1474.1; -

CC EMBL, X07024; CAA30073.1; ALT_SEQ.

CC PIR; A40262; A40262.

CC PDB; 1ROF; 07-JUN-00.

CC TRANSFAC; T02206; -

CC Genew; HGNC:11535; TAF1.

CC MIM; 313650; -

CC DR GO; GO:0005669; C:transcription factor TFIID complex; TAS.

CC DR GO; GO:0016251; F:general RNA polymerase II transcription fac. . .; TAS.

CC DR GO; GO:0004672; F:protein kinase activity; TAS.

CC DR GO; GO:0000114; F:G1-specific transcription in mitotic cell c. . .; TAS.

CC DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.

CC DR GO; GO:0006461; P:protein complex assembly; TAS.

CC InterPro; IPR001487; Bromodomain.

DR Pfam; PF00439; bromodomain; 2.
 DR PRINTS; PR00503; BROMODOMAIN.
 DR SMART; SM00297; BROMO; 2.
 DR PROSITE; PS00633; BROMODOMAIN 1; 2.
 DR PROSITE; PS00614; BROMODOMAIN 2; 2.
 KM Bromodomain; Nuclear protein; DNA-binding; Cell cycle; Repeat;
 KM Transcription regulation; Phosphorylation; 3D-structure.
 FT DOMAIN 157 165 PRO-RICH.
 FT DNA BIND 1195 1273 HMG BOX (POTENTIAL).
 FT DOMAIN 1351 1358 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 1387 1467 BROMODOMAIN 1.
 FT DOMAIN 1520 1590 BROMODOMAIN 2.
 FT DOMAIN 1627 1872 ASP/GLU-RICH (ACIDIC TAIL).
 SQ SEQUENCE 1872 AA; 212676 MW; 93BB3D181A72ABEB CRC64;

Query Match 36.4%; Score 40; DB 1; Length 1872;
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

4 PNKGTQGYTDQIERPL 19
 ||||| : : :
 969 PNKGTQGYTDQIERPLV 984

RESULT 32
 RS9_MYCTU STANDARD; PRT; 151 AA.
 AC 006359;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S9.
 GN RPS1 OR RV3442C OR MT3547 OR MTCY7.14C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=96295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers R.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Rutherford J.E., Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RA complete genome sequence."
 RL Nature 393:537-544 (1998).
 RU [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Dvorkin T., Feldman D., Kouri H., Gill J., Mikula A.,
 RA Bhat N.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RA laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL; Z95389; CAB08691.1; -.
 DR EMBL; AE007159; AAK47888.1; -.
 DR PIR; H70976; H70976.
 DR TIGR; MT3547; -.
 DR TUBERCULIN; RV3442C; -.
 DR HAMAP; MF_00532; -; 1.
 DR InterPro; IPR000754; Ribosomal_S9.
 DR Pfam; PF00380; Ribosomal_S9; 1.
 DR ProDom; PD001627; Ribosomal_S9; 1.
 DR PROSITE; PS00360; RIBOSOMAL_S9; 1.
 KM Ribosomal protein; Complete proteome.
 SQ SEQUENCE 151 AA; 16436 MW; 674679A0A1C2FC9E CRC64;

Query Match 35.9%; Score 39.5; DB 1; Length 151;
 Best Local Similarity 52.6%; Pred. No. 20;
 Matches 10; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 2 YFPNKGTYQGYTDQIERPLM 20
 ||||| : : :
 DB 59 YFPNKGTYQGYTDQIERPLV 74

RESULT 33
 RS9_MYCTU STANDARD; PRT; 153 AA.
 AC P40828;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S9.
 GN RPS1 OR M0365 OR B229_C2_191.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RP SEQUENCE FROM N.A.
 RC Smith D.R., Robison K.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RU [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21126732; PubMed=11234002;
 RA Cole S.T., Eigemeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Dutfoy S., Felwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrett B.G.;
 RA "Massive gene decay in the leprosy bacillus."
 RL Nature 409:1007-1011 (2001).
 RU [2]
 RP SEQUENCE FROM N.A.
 CC -1- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U00020; AAI1296.1; -.
 CC EMBL; AL583918; CAC29873.1; -.
 CC PIR; S72982; S72982.
 CC Leproma; ML0365; -.
 CC HAMAP; MF_00532; -; 1.
 CC InterPro; IPR000754; Ribosomal_S9.
 CC Pfam; PF00380; Ribosomal_S9; 1.
 CC ProDom; PD001627; Ribosomal_S9; 1.
 CC PROSITE; PS00360; RIBOSOMAL_S9; 1.

KW Ribosomal protein; Complete proteome.
 RX SEQUENCE 153 AA; 16911 MW; BBPD7CD1029F0735 CRC64;

Query Match 35.5%; Score 39.5; DB 1; Length 153;
 Best Local Similarity 52.6%; Pred. No. 20;
 Matches 10; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 2 TYPENKGTQYTDQIERPLM 20
 ||||| : |||
 Db 61 TYPENKVTQQL--IKAPLV 76

RESULT 34

HB2 CATMA STANDARD; PRT; 141 AA.
 AC P82112;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 RT Hemoglobin alpha-2 chain (HB 2).
 CC Catharacta macconnicki (South polar skua).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Charadriiformes; Stercorariidae;
 CC Catharacta.
 CC NCBI_TaxID=94055;
 RX SEQUENCE.
 RC TISSUE=Blood;
 RX MEDLINE=20453132; PubMed=10998071;
 RA Tamburini M., Riccio A., Romano M., Giardina B., di Prisco G.;
 RT "Structural and functional analysis of the two hemoglobins of the
 RT antarctic seabird Catharacta macconnicki. Characterization of an
 RT additional phosphate binding site by molecular modelling.";
 RL Eur. J. Biochem. 267:6089-6098(2000).
 CC -1- FUNCTION: Involved in oxygen transport from the lung to the
 CC various peripheral tissues.
 CC -1- SUBUNIT: Heterotrimer of two alpha chains and two beta chains.
 CC -1- TISSUE SPECIFICITY: Red blood cells.
 CC -1- SIMILARITY: Belongs to the globin family.
 DR HSSP: P02001; IHR.
 DR InterPro: IPR002338; Alpha_haem.
 DR InterPro: IPR000971; Globin.
 DR Pfam: PF00042; globin; 1.
 DR PRINTS: PR00612; ALPHAHAE.
 DR PROSITE: PS01033; GLOBIN; 1.
 KM Heme; Oxygen transport; Transport; Erythrocyte.
 FT METAL 58 87 IRON (HEME DISTAL LIGAND).
 FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
 SEQUENCE 141 AA; 15930 MW; F3535256589083C4 CRC64;

Query Match 35.5%; Score 39; DB 1; Length 141;
 Best Local Similarity 53.3%; Pred. No. 23;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYPENKGTQYTDQI 15
 ||||| : |||
 Db 41 TYPFHDLDGSDQI 55

RESULT 35
 HBAD PASMO STANDARD; PRT; 141 AA.
 AC P07413;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hemoglobin alpha-D chain.
 GN HBAD.

OS Passer montanus (Tree sparrow).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Passeriformes; Passeridae;
 CC Passer.
 CC NCBI_TaxID=9160;

RN [1]
 RP SEQUENCE.
 RX MEDLINE=86077293; PubMed=4074502;
 RA Schlegel D., Braunitzer G., Oberthur W., Koesters J., Grimm F.;
 RT "Hemoglobin of tree sparrows (Passer montanus, Passeriformes):
 RT Sequence of the major (Hb A) and minor (Hb D) components.";
 RL Biol. Chem. Hoppe-Seyler 366:893-899(1985).
 CC -1- FUNCTION: Involved in oxygen transport from the lung to the
 CC various peripheral tissues.
 CC -1- SUBUNIT: Heterotrimer of two alpha-D chains and two beta chains.
 CC -1- TISSUE SPECIFICITY: Red blood cells.
 CC -1- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor
 CC hemoglobin component, called hemoglobin d, which is expressed in
 CC late embryonic and adult life.
 CC -1- SIMILARITY: Belongs to the globin family.
 DR HSSP: P02001; IHR.
 DR InterPro: IPR002338; Alpha_haem.
 DR InterPro: IPR000971; Globin.
 DR Pfam: PF00042; globin; 1.
 DR PRINTS: PR00612; ALPHAHAE.
 DR PROSITE: PS01033; GLOBIN; 1.
 KM Heme; Oxygen transport; Transport; Erythrocyte.
 FT METAL 58 87 IRON (HEME DISTAL LIGAND).
 FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
 SEQUENCE 141 AA; 15728 MW; BB6220406F9E0B90 CRC64;

Query Match 35.5%; Score 39; DB 1; Length 141;
 Best Local Similarity 53.3%; Pred. No. 23;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYPENKGTQYTDQI 15
 ||||| : |||
 Db 41 TYPFHDLDGSDQI 55

RESULT 36

HBAD STUVU STANDARD; PRT; 141 AA.
 AC P02004;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hemoglobin alpha-D chain.
 GN HBAD.
 OS Sturnus vulgaris (Starling).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Passeriformes; Sturnidae; Sturnus.
 CC NCBI_TaxID=9172;
 RX SEQUENCE.
 RX MEDLINE=84184202; PubMed=6714943;
 RA Oberthur W., Braunitzer G.;
 RT "Hemoglobins of the common starling (Sturnus vulgaris, Passeriformes).
 RT The primary structures of the alpha, alpha-d and beta chains.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 365:159-173(1984).
 CC -1- FUNCTION: Involved in oxygen transport from the lung to the
 CC various peripheral tissues.
 CC -1- SUBUNIT: Heterotrimer of two alpha-D chains and two beta chains.
 CC -1- TISSUE SPECIFICITY: Red blood cells.
 CC -1- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor
 CC hemoglobin component, called hemoglobin d, which is expressed in
 CC late embryonic and adult life.
 CC -1- SIMILARITY: Belongs to the globin family.
 DR PIR: A02331; HA5B.
 DR HSSP: P02001; IHR.
 DR InterPro: IPR002338; Alpha_haem.
 DR InterPro: IPR000971; Globin.
 DR Pfam: PF00042; globin; 1.
 DR PRINTS: PR00612; ALPHAHAE.
 DR PROSITE: PS01033; GLOBIN; 1.
 KM Heme; Oxygen transport; Transport; Erythrocyte.
 FT METAL 58 58 IRON (HEME DISTAL LIGAND).

FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
 SQ SEQUENCE 141 AA; 15660 MW; 647C64086BC97D89 CRC64;
 Query Match 35.5%; Score 39; DB 1; Length 141;
 Best Local Similarity 53.3%; Pred. No. 23;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQI 15
 |||||
 41 TYFPNKGTOQYTDQI 55

Db

RESULT 37
 IE0 NPVOIP STANDARD; PRT; 245 AA.
 AC 010369;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DT Immediate-early protein IE-0.
 IE-0.
 Oryzia pseudotsugata multicapsid polyhedrosis virus (OPMVV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OK NCBI_TaxID=164623;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=97271300; PubMed=9126251;
 RA Athens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
 RA Rohmann G.F.;
 RT "The sequence of the Oryzia pseudotsugata multinucleocapsid nuclear
 RT polyhedrosis virus genome.";
 RL Virology 229:381-399(1997).
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U75930; AAC59137.1; -;
 DR InterPro: IPR001841; Znf_ring.
 DR Pfam: PF05290; Baculo_IE-1; 1.
 DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
 DR PROSITE: PS50089; ZF_RING_2; 1.
 DR Early protein; Zinc-finger.
 ZN FING 195 240
 SQ SEQUENCE 245 AA; 27117 MW; 357FA3B15F7B1029 CRC64;

QY 1 TYFPNKGTOQYTDQI 16
 |||||
 5 THWPLVLSKSYTDACE 20

Db

RESULT 38
 VS09 ROTH9 STANDARD; PRT; 326 AA.
 AC P33492;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glycoprotein VP7 (Serotype-specific antigen) (outer shell
 DE glycoprotein).
 GN S9.
 OS Human rotavirus (serotype G / strain RK9).
 OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.

OK NCBI_TaxID=33722;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91214432; PubMed=1965285;
 RA Bessard I.N., Novikova P.A., Borodin A.M.;
 RT "Use of the polymerase chain reaction for analysis of rotaviruses.
 RT Nucleotide sequence of a gene, coding for the basic neutralizing
 RT antigen VP7 of a human rotavirus with a new G-serotype (letter).";
 RL Bioorg. Khim. 16:1689-1691(1990).
 CC -1- FUNCTION: VP7 IS THE MAJOR NEUTRALIZATION ANTIGEN.
 CC -1- SUBCELLULAR LOCATION: Rough endoplasmic reticulum.
 CC PIR: JN0307; JN0307.
 DR InterPro: IPR001963; VP7.
 DR Pfam: PF00434; VP7; 1.
 DR Prodom: PD000191; VP7; 1.
 KM Coat protein; Transmembrane; Glycoprotein.
 FT TRANSMEM 32 48
 FT CARBOHYD 69 69
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 326 AA; 37323 MW; 3C2439114E23A8FE CRC64;

QY 2 YFPNKGTOQYTD 13
 |||||
 84 YFPNDVNTQIRD 95

Db

Query Match 35.5%; Score 39; DB 1; Length 326;
 Best Local Similarity 58.3%; Pred. No. 57;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RESULT 39
 ID SPEB_STRPY STANDARD; PRT; 398 AA.
 AC P00788; P26296; Q54960; Q54962; Q54963; Q54964; Q54965;
 AC Q54966; Q54967; Q54968; Q57024; Q57082; Q57202; Q57211; Q57212;
 AC Q55680;
 DT 21-JUN-1986 (Rel. 01, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Streptococcal precursor (EC 3.4.22.10) (Streptococcal cysteine
 DE protease) (Streptococcus peptidase A) (SPP) (Exotoxin type B) (SPE
 DE B).
 GN SPEB OR SPY2039 OR SPYM3_1742 OR SPS1739 OR SPYM18_2099.
 OS Streptococcus pyogenes,
 OS Streptococcus pyogenes (serotype M3), and
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OK NCBI_TaxID=1314, 198466, 186103;
 [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 28-32 AND 146-162.
 RP STRAIN=86-858, and NY-5;
 RX MEDLINE=90330563; PubMed=2198264;
 RA Hausser A.R., Schlevert P.M.;
 RT "Nucleotide sequence of the streptococcal pyrogenic exotoxin type B
 RT gene and relationship between the toxin and the streptococcal
 RT proteinase precursor".
 RL U. Bacteriol. 172:4536-4542(1990).
 [2]
 RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RP STRAIN=789 / Serotype M1, 327 / Serotype M2, 75 / Serotype M3,
 RC 157 / Serotype M3, 315 / Serotype M3, 1251 / Serotype M3,
 RC 321 / Serotype M4, 1289 / Serotype M5, 303 / Serotype M6,
 RC 429 / Serotype M8, 796 / Serotype M9, 800 / Serotype M9,
 RC 1896 / Serotype M10, 650 / Serotype M11, 282 / Serotype M12,
 RC 653 / Serotype M13, 660 / Serotype M14, 1898 / Serotype M15,
 RC 1233 / Serotype M17, 156 / Serotype M18, 300 / Serotype M18,
 RC 1294 / Serotype M19, 162 / Serotype M22, 1901 / Serotype M23,
 RC 684 / Serotype M24, 686 / Serotype M25, 1838 / Serotype M27,
 RC 587 / Serotype M28, 366 / Serotype M30, 427 / Serotype M31,
 RC 807 / Serotype M33, 1841 / Serotype M41, 1842 / Serotype M43,
 RC 1226 / Serotype M44, 719 / Serotype M49, 1864 / Serotype M43,
 RC 1882 / Serotype M59, 168 / Serotype M66, 302 / Serotype M73,
 RC 758 / Serotype M75, 1911 / Serotype M75, 1832 / Serotype M76,

RC 165 / Serotype M, 262 / Serotype M, 317 / Serotype M,
RC 1990 / Serotype M, 1991 / Serotype M, 2017 / Serotype M,
RC 1990 / Serotype M, 1719 / Serotype T8, 269 / Serotype T28,
RC 1590, 1870, 1871, 1872, 1893, and 1914A;
RC MEDLINE=94285789; PubMed=7516997;
RA Kapur V., Topouzis S., Majeky M.W., Li L.L., Hamrick M.R.,
RA Hamill R.J., Patti J.M., Musser J.M.;
RT "A conserved Streptococcus pyogenes extracellular cysteine protease
RT cleaves human fibronectin and degrades vitronectin.";
RL Microb. Pathog. 15:327-346(1993).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=A-20 / Serotype M1,T1;
RA Wu J.-J.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN=SV / Serotype M23;
RA Hong K.;
RL "A novel cloning method used arbitrarily primed PCR.";
RN Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN (5)
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Szate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.F.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus
RT pyogenes.";
RN Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN (6)
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phase-encoded toxins, the high-virulence phenotype and clone
RT emergence.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN (7)
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RX Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS8232.";
RN Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN (8)
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Campel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.W., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
RN (9)
RP PRELIMINARY SEQUENCE OF 28-86 AND 121-398.
RA Yonaha K., Elliott S.D., Liu T.-Y.;
RT "Primary structure of zymogen of streptococcal proteinase.";
RN J. Protein Chem. 1:317-334(1982).
RN (10)
RP PRELIMINARY SEQUENCE OF 146-398.
RA MEDLINE=76190087; PubMed=1270417;

RA Tai J.Y., Kortt A.A., Liu T.-Y., Elliott S.D.;
RT "Primary structure of streptococcal proteinase. III. Isolation of
RT cyanogen bromide peptides: complete covalent structure of the
RT polypeptide chain.";
RN J. Biol. Chem. 251:1955-1959(1976).
RN (11)
RP FUNCTION.
RC STRAIN=N2131 / Serotype M49,T14;
RX MEDLINE=99081733; PubMed=9864206;
RA Kuo C.-F., Wu J.-J., Tsai P.-J., Kao F.-J., Lei H.-Y., Lin M.T.,
RA Lin Y.-S.;
RT "Streptococcal pyrogenic exotoxin B induces apoptosis and reduces
RT phagocytic activity in U937 cells.";
RL Infect. Immun. 67:126-130(1999).
RN (12)
RP FUNCTION.
RC STRAIN=A-20 / Serotype M1,T1;
RX MEDLINE=99386817; PubMed=10456871;
RA Tsai P.-J., Lin Y.-S., Kuo C.-F., Lei H.-Y., Wu J.-J.;
RT "Group A Streptococcus induces apoptosis in human epithelial cells.";
RL Infect. Immun. 67:4334-4339(1999).
RN (13)
RP FUNCTION: IMPORTANT STREPTOCOCCAL VIRULENCE FACTOR WHICH CLEAVES
CC HUMAN FIBRONECTIN AND DEGRADES VITRONECTIN. ALSO CLEAVES HUMAN
CC IL1B PRECURSOR TO FORM BIOLOGICALLY ACTIVE IL1B. CAN INDUCE
CC APOPTOSIS IN HUMAN MONOCYTES AND EPITHELIAL CELLS IN VITRO, AND
CC REDUCES PHAGOCYTIC ACTIVITY IN MONOCYTIC CELLS. THUS, MAY PLAY A
CC ROLE IN BACTERIAL COLONIZATION, INVASION, AND INHIBITION OF WOUND
CC HEALING.
CC
CC -1- CATALYTIC ACTIVITY: Preferential cleavage with hydrophobic
CC residues at P2, P1 and P1'.
CC
CC -1- SUBCELLULAR LOCATION: Secreted.
CC
CC -1- DISEASE: THE STREPTOCOCCAL PYROGENIC TOXINS A, B, AND C ARE
CC THE CAUSATIVE AGENTS OF THE SYMPTOMS ASSOCIATED WITH SCARLET
CC FEVER. HAVE BEEN ASSOCIATED WITH STREPTOCOCCAL TOXIC SHOCK-LIKE
CC DISEASE AND MAY PLAY A ROLE IN THE EARLY EVENTS OF RHEUMATIC
CC FEVER.
CC
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C10.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
CC or send an email to license@sib.ch).
CC
CC -----
CC EMBL: M86905; AAA26978.1; -
CC EMBL: L26125; AAA26979.1; -
CC EMBL: L26126; AAA26992.1; -
CC EMBL: L26127; AAA26993.1; -
CC EMBL: L26128; AAA26994.1; -
CC EMBL: L26130; AAA26996.1; -
CC EMBL: L26131; AAA26997.1; -
CC EMBL: L26132; AAA26998.1; -
CC EMBL: L26133; AAA26999.1; -
CC EMBL: L26134; AAA27000.1; -
CC EMBL: L26135; AAA27001.1; -
CC EMBL: L26136; AAA27002.1; -
CC EMBL: L26137; AAA27003.1; -
CC EMBL: L26138; AAA27004.1; -
CC EMBL: L26139; AAA27005.1; -
CC EMBL: L26140; AAA27006.1; -
CC EMBL: L26141; AAA27007.1; -
CC EMBL: L26142; AAA27008.1; -
CC EMBL: L26143; AAA27009.1; -
CC EMBL: L26144; AAA27010.1; -
CC EMBL: L26145; AAA27011.1; -
CC EMBL: L26146; AAA27012.1; -
CC EMBL: L26147; AAA27013.1; -
CC EMBL: L26148; AAA27014.1; -
CC EMBL: L26149; AAA27015.1; -
CC EMBL: L26150; AAA27016.1; -
CC EMBL: L26151; AAA26980.1; -

```

DR EMBL; L26152; AAA26981.1; -
DR EMBL; L26153; AAA26982.1; -
DR EMBL; L26154; AAA26983.1; -
DR EMBL; L26155; AAA26984.1; -
DR EMBL; L26156; AAA26985.1; -
DR EMBL; L26157; AAA26986.1; -
DR EMBL; L26158; AAA26987.1; -
DR EMBL; L26159; AAA26988.1; -
DR EMBL; L26160; AAA26989.1; -
DR EMBL; L26161; AAA26990.1; -
DR EMBL; L26162; AAA26991.1; -
DR EMBL; AF104940; AAD17930.1; -

Query Match      35.5%; Score 39; DB 1; Length 398;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 PPNKGTQOYT 12
    :|||:|
    206 YPNKGLKDYT 215

RESULT 40
ID YAFB_ECOLI STANDARD; PRT; 414 AA.
AC P04335; P77413;
DT 20-MAR-1987 (Rel. 04, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yafA.
GN YAFB OR B0239.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85155481; PubMed=6397401;
RA Nuesch J., Schumperli D.;
RT "Structural and functional organization of the gpt gene region of
RL Escherichia coli."
RL Gene 32:243-249(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RL 4.0 - 6.0 min (189,987 - 281,416bp) region."
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
Davis K., Federspiel N., Hyman R., Kallman S., Komp C., Kurd O.,
Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13422; AAA23929.1; -
DR EMBL; AB000132; AAC73343.1; -
DR EMBL; D83536; BAA77908.1; -
DR EMBL; U70214; AAB06659.1; -
DR PIR; H64748; O08C49.
DR EcoGene; EG11091; yafA.
DR InterPro; IPR00379; Ser_estra_site.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 242 242 T -> S (IN REF. 1 AND 3).
FT CONFLICT 277 278 NV -> TL (IN REF. 1 AND 3).
FT CONFLICT 397 414 NFDKGLQETDWEIKRLC -> KLTIVFRKSPGSKNACV
FT KNLNLFANLYKQHLHNR (IN REF. 1 AND 3).
SQ SEQUENCE 414 AA; 47008 MW; 2C154E46FC9565E CRC64;

Query Match      35.5%; Score 39; DB 1; Length 414;
Best Local Similarity 53.3%; Pred. No. 74;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 NKGTQYTDTERPL 19
    :|||:|
    399 DKGLQETDWEIKRL 413

Search completed: August 20, 2003, 09:30:49
Job time : 13.564 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 58.1443 Seconds
(without alignments)
88.763 Million cell updates/sec

Title: US-09-991-433-48
Perfect score: 110
Sequence: 1 TYPFNKGTQYTDIERPLM 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacterioph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	110	100.0	138	12	Q85155	Q85155 human parvo
2	110	100.0	141	12	Q85171	Q85171 human parvo
3	110	100.0	141	12	Q85166	Q85166 human parvo
4	110	100.0	141	12	Q85146	Q85146 human parvo
5	110	100.0	142	12	Q85168	Q85168 human parvo
6	110	100.0	142	12	Q85173	Q85173 human parvo
7	110	100.0	144	12	Q85138	Q85138 human parvo
8	110	100.0	145	12	Q85181	Q85181 human parvo
9	110	100.0	145	12	Q85161	Q85161 human parvo
10	110	100.0	146	12	Q85158	Q85158 human parvo
11	110	100.0	147	12	Q85142	Q85142 human parvo
12	110	100.0	148	12	Q85150	Q85150 human parvo
13	110	100.0	151	12	Q85131	Q85131 human parvo
14	110	100.0	151	12	Q85177	Q85177 human parvo
15	110	100.0	153	12	Q85123	Q85123 human parvo
16	110	100.0	157	12	Q85196	Q85196 human parvo

17	110	100.0	162	12	Q85135	Q85135 human parvo
18	110	100.0	546	12	Q913X0	Q913X0 human parvo
19	110	100.0	546	12	Q913W7	Q913W7 human parvo
20	110	100.0	554	12	Q9PZS9	Q9PZS9 human parvo
21	110	100.0	554	12	Q90201	Q90201 human parvo
22	110	100.0	554	12	Q65790	Q65790 human parvo
23	110	100.0	554	12	Q912B7	Q912B7 human eryth
24	110	100.0	554	12	Q9JGP7	Q9JGP7 human parvo
25	110	100.0	554	12	Q9WKL9	Q9WKL9 human parvo
26	110	100.0	554	12	Q8JMS4	Q8JMS4 human parvo
27	110	100.0	760	12	Q9PZT8	Q9PZT8 human parvo
28	110	100.0	761	12	Q9PZU0	Q9PZU0 human parvo
29	110	100.0	765	12	Q9PZT6	Q9PZT6 human parvo
30	110	100.0	769	12	Q9PZT4	Q9PZT4 human parvo
31	110	100.0	773	12	Q913X1	Q913X1 human parvo
32	110	100.0	773	12	Q913W8	Q913W8 human parvo
33	110	100.0	781	12	Q8JMS6	Q8JMS6 human parvo
34	110	100.0	781	12	Q65789	Q65789 human parvo
35	110	100.0	781	12	P90223	P90223 human parvo
36	110	100.0	781	12	Q85191	Q85191 human parvo
37	110	100.0	781	12	P90221	P90221 human parvo
38	110	100.0	781	12	P89318	P89318 human parvo
39	110	100.0	781	12	Q8JTE3	Q8JTE3 erythrocyte
40	110	100.0	781	12	P90224	P90224 human parvo
41	110	100.0	781	12	Q9PZT0	Q9PZT0 human parvo
42	110	100.0	781	12	P89316	P89316 human parvo
43	110	100.0	781	12	P89320	P89320 human parvo
44	110	100.0	781	12	Q85117	Q85117 human parvo
45	110	100.0	781	12	P89321	P89321 human parvo

ALIGNMENTS

RESULT 1

ID Q85155 PRELIMINARY; PRT; 138 AA.
AC Q85155;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10796;
RX [1]
RP SEQUENCE FROM N.A.
RA Hemauey A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G., Wolf H., Modrow S.;
RT "XXXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z70564; CA94477.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15237 MW; 5899FB8879A3E6BD CRC64;

Query Match 100.0%; Score 110; DB 12; Length 138;
Best Local Similarity 100.0%; Pred. No. 7.6e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TYPFNKGTQYTDIERPLM 20
DB 103 TYPFNKGTQYTDIERPLM 122

RESULT 2

ID Q85171 PRELIMINARY; PRT; 141 AA.
AC Q85171;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z70580; CA94493.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 141
SQ SEQUENCE 141 AA; 15770 MW; C9E92572A78C6C27 CRC64;
Query Match 100.0%; Score 110; DB 12; Length 141;
Best Local Similarity 100.0%; Pred. No. 7.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOYTDQIERPLM 20
Db 105 TYFPNKGTOYTDQIERPLM 124
RESULT 3
ID 085166 PRELIMINARY; PRT; 141 AA.
AC 085166;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z70575; CA94488.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 141
SQ SEQUENCE 141 AA; 15765 MW; 4115D3D915751757 CRC64;
Query Match 100.0%; Score 110; DB 12; Length 141;
Best Local Similarity 100.0%; Pred. No. 7.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOYTDQIERPLM 20
Db 105 TYFPNKGTOYTDQIERPLM 124
RESULT 4
ID 085146 PRELIMINARY; PRT; 141 AA.
AC 085146;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VPI.
OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z70555; CA94467.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 141
SQ SEQUENCE 141 AA; 15712 MW; 32F6B23E7E6B583 CRC64;
Query Match 100.0%; Score 110; DB 12; Length 141;
Best Local Similarity 100.0%; Pred. No. 7.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOYTDQIERPLM 20
Db 102 TYFPNKGTOYTDQIERPLM 121
RESULT 5
ID 085168 PRELIMINARY; PRT; 142 AA.
AC 085168;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z70577; CA94490.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 142
SQ SEQUENCE 142 AA; 15821 MW; 2B6E4D9A6784F8C4 CRC64;
Query Match 100.0%; Score 110; DB 12; Length 142;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOYTDQIERPLM 20
Db 101 TYFPNKGTOYTDQIERPLM 120
RESULT 6
ID 085173 PRELIMINARY; PRT; 142 AA.
AC 085173;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,

RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270562; CA94495.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 142
 SQ SEQUENCE 142 AA; 15820 MW; 3A95E07C0BC0434 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 142;
 Best Local Similarity 100.0%; Pred. No. 7.9e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPFNKGTQOYTDQIERPLM 20
 DB 101 TYPFNKGTQOYTDQIERPLM 120

RESULT 7

AC 085138 PRELIMINARY; PRT; 144 AA.
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemenauer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270547; CA94459.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 144
 SQ SEQUENCE 144 AA; 16022 MW; 95D6EF37BA5997AB CRC64;

Query Match 100.0%; Score 110; DB 12; Length 144;
 Best Local Similarity 100.0%; Pred. No. 8e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYPFNKGTQOYTDQIERPLM 20
 103 TYPFNKGTQOYTDQIERPLM 122

RESULT 8

ID 085181 PRELIMINARY; PRT; 145 AA.
 AC 085181;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemenauer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270590; CA94503.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16210 MW; 6B45A1E1B9E923C4 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 8.1e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPFNKGTQOYTDQIERPLM 20
 DB 105 TYPFNKGTQOYTDQIERPLM 124

RESULT 9

ID 085161 PRELIMINARY; PRT; 145 AA.
 AC 085161;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemenauer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270570; CA94483.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16165 MW; B4FE95AB6E4D9A67 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 8.1e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPFNKGTQOYTDQIERPLM 20
 DB 101 TYPFNKGTQOYTDQIERPLM 120

RESULT 10

ID 085158 PRELIMINARY; PRT; 146 AA.
 AC 085158;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemenauer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270567; CA94480.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 146
 SQ SEQUENCE 146 AA; 16252 MW; 2B74FE95AB6E4D9A CRC64;

Query Match 100.0%; Score 110; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 8.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 101 TYPFNKGTQOYTDQIERPLM 120

RESULT 11

085142 PRELIMINARY; PRT; 147 AA.

AC 085142; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GN Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giggler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates."
DL EMBL; 270551; CA94463.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 147 147

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 104 TYPFNKGTQOYTDQIERPLM 123

Query Match 100.0%; Score 110; DB 12; Length 147;
Best Local Similarity 100.0%; Pred. No. 8.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 104 TYPFNKGTQOYTDQIERPLM 123

RESULT 12

085150 PRELIMINARY; PRT; 148 AA.

AC 085150; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GN Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

RA Hemauer A., Von Poblotszki A., Giggler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates."
DL EMBL; 270559; CA94471.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 148 148

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 105 TYPFNKGTQOYTDQIERPLM 124

Query Match 100.0%; Score 110; DB 12; Length 148;
Best Local Similarity 100.0%; Pred. No. 8.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 105 TYPFNKGTQOYTDQIERPLM 124

DB 104 TYPFNKGTQOYTDQIERPLM 123

RESULT 13

085131 PRELIMINARY; PRT; 151 AA.

AC 085131; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GN Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giggler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates."
DL EMBL; 270540; CA94452.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 151 151

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 105 TYPFNKGTQOYTDQIERPLM 124

Query Match 100.0%; Score 110; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 105 TYPFNKGTQOYTDQIERPLM 124

RESULT 14

085177 PRELIMINARY; PRT; 151 AA.

AC 085177; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GN Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

RA Hemauer A., Von Poblotszki A., Giggler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates."
DL EMBL; 270586; CA94499.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 151 151

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 105 TYPFNKGTQOYTDQIERPLM 124

Query Match 100.0%; Score 110; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 105 TYPFNKGTQOYTDQIERPLM 124

RESULT 15

```

085123
ID 085123 PRELIMINARY; PRT; 153 AA.
AC 085123;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96332516; PubMed=8760426;
RA Hemanter A., Von Pobioltzki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
  "Sequence variability among different parvovirus B19 isolates.";
  J. Gen. Virol. 77:1781-1785(1996).
RL EMBL; Z70532; CA94444.1; -.
DR InterPro; IPR001403; Parvo_coat.
PFam; PF00740; Parvo_coat; 1.
FT NON_TER 1 153
SQ SEQUENCE 153 AA; 17331 MW; ECCA1F44020814EC CRC64;

Query Match 100.0%; Score 110; DB 12; Length 153;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOYTDQIERPLM 20
Db 104 TYFPNKGTOYTDQIERPLM 123

RESULT 16
085196 PRELIMINARY; PRT; 157 AA.
AC 085196;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RX Hemanter A., Von Pobioltzki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
  "XXXSequence variability among different parvovirus B19 isolates.";
  Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70603; CA94518.1; -.
DR InterPro; IPR001403; Parvo_coat.
PFam; PF00740; Parvo_coat; 1.
FT NON_TER 1 157
SQ SEQUENCE 157 AA; 17591 MW; 1D42191887FFCE03 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 157;
Best Local Similarity 100.0%; Pred. No. 8.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOYTDQIERPLM 20
Db 106 TYFPNKGTOYTDQIERPLM 125

RESULT 17
085135 PRELIMINARY; PRT; 162 AA.
AC 085135;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

```

```

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemanter A., Von Pobioltzki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
  "XXXSequence variability among different parvovirus B19 isolates.";
  Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70544; CA94456.1; -.
DR InterPro; IPR001403; Parvo_coat.
PFam; PF00740; Parvo_coat; 1.
FT NON_TER 1 162
SQ SEQUENCE 162 AA; 18005 MW; 3F51443566660P2 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 162;
Best Local Similarity 100.0%; Pred. No. 9.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOYTDQIERPLM 20
Db 105 TYFPNKGTOYTDQIERPLM 124

RESULT 18
0913X0 PRELIMINARY; PRT; 546 AA.
AC 0913X0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DE 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major virus capsid protein VP2 (Fragment).
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lali;
RA Hokyvar K., Soderlund-Venemo M., Ranki A., Kiviluoto O., Partio E.K.,
  Hedman K.;
  "A new parvovirus B19 genotype persistent in skin.";
  Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY044266; AAK95572.1; -.
DR InterPro; IPR001403; Parvo_coat.
PFam; PF00740; Parvo_coat; 1.
FT NON_TER 546 546
SQ SEQUENCE 546 AA; 59934 MW; 467BB468A67282B5 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 3.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOYTDQIERPLM 20
Db 391 TYFPNKGTOYTDQIERPLM 410

RESULT 19
0913W7 PRELIMINARY; PRT; 546 AA.
AC 0913W7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DE 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major virus capsid protein VP2 (Fragment).
GN VP2.
OS Human parvovirus B19.

```

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HAM;
 RA Hokynar K., Soderlund-Venermo M., Ranki A., Kiviluoto O., Partio E.K., Hedman K.;
 RT "A new parvovirus B19 genotype persistent in skin."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY044268; AAK95575.1; -
 DR InterPro: IPR001403; Parvo_coat;
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON TER 546 546
 SQ SEQUENCE 546 AA; 59910 MW; BBPIID23695E3FCD CRC64;

Query Match 100.0%; Score 110; DB 12; Length 546;
 Best Local Similarity 100.0%; Pred. No. 3.5e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYPENKGTQOYTDOIERPLM 20
 |||||
 391 TYPENKGTQOYTDOIERPLM 410

RESULT 20
 O9P2S9 PRELIMINARY; PRT; 554 AA.
 AC O9P2S9;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE VP2 capsid protein.
 GN VP.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HV;
 RA Gallinella G., Venturoli S.;
 RT "B19 Genome Sequence and Structure Analysis."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF162273; AAD4615.1; -
 DR InterPro: IPR001403; Parvo_coat;
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60833 MW; 7901FEBD65A97E6 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 3.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYPENKGTQOYTDOIERPLM 20
 |||||
 391 TYPENKGTQOYTDOIERPLM 410

RESULT 21
 O90201 PRELIMINARY; PRT; 554 AA.
 ID O90201;
 AC O90201;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE VP2 structural protein (Capsid protein VP2) (Frigment).
 GN VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97081188; PubMed=8922470;
 RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the

RT VP1/VP2 gene from multiple isolates.";
 RL J. Gen. Virol. 77:2767-2774(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WT, and N8;
 RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K., Sasaki T.;
 RT "Sequence of human parvovirus B19 isolates from patients with rheumatoid arthritis."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N8;
 RX MEDLINE=90218047; PubMed=2157807;
 RA Umene K., Nunoue T.;
 RT "The genome type of human parvovirus B19 strains isolated in Japan during 1981 differs from types detected in 1986 to 1987: a correlation between genome type and prevalence.";
 RL J. Gen. Virol. 71:983-986(1990).
 DR EMBL: U53595; AAB47453.1; -
 DR EMBL: U53596; AAB47455.1; -
 DR EMBL: U53593; AAB47449.1; -
 DR EMBL: U53594; AAB47451.1; -
 DR EMBL: U53597; AAB47457.1; -
 DR EMBL: U53600; AAB47463.1; -
 DR EMBL: U53601; AAB47459.1; -
 DR EMBL: U53598; AAB47459.1; -
 DR EMBL: U53599; AAB47461.1; -
 DR EMBL: AB030693; BAA90290.1; -
 DR EMBL: AB030673; BAA90268.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON TER 554 554
 SQ SEQUENCE 554 AA; 60853 MW; 98FE598FE20CB66F CRC64;

Query Match 100.0%; Score 110; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 3.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYPENKGTQOYTDOIERPLM 20
 |||||
 391 TYPENKGTQOYTDOIERPLM 410

RESULT 22
 O65790 PRELIMINARY; PRT; 554 AA.
 ID O65790;
 AC O65790;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VENI;
 RA Echavarría Mayo J.E., Erdman D.D.;
 RT Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U1358; AAB3559.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60798 MW; BA89F2B23B54E24 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 3.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYPENKGTQOYTDOIERPLM 20
 |||||
 391 TYPENKGTQOYTDOIERPLM 410

```

RESULT 23
ID Q912B7 PRELIMINARY; PRT; 554 AA.
AC Q912B7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE VP2 protein.
GN VP2.
OS Human erythrovirus V9.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=72197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V9;
RA Nguyen O.T.;
RT "Molecular cloning and sequencing of a novel human erythrovirus genome: new species beside B19 in the genus Erythrovirus.";
Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AJ249437; CAC80622.1; -.
InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat.1.
SQ SEQUENCE 554 AA; 60885 MW; B4F38CDA80F336 CRC64;

Query Match
Best Local Similarity 100.0%; Score 110; DB 12; Length 554;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 391 TYPFNKGTQOYTDQIERPLM 410

RESULT 24
ID Q9UGP7 PRELIMINARY; PRT; 554 AA.
AC Q9UGP7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Capsid protein VP2.
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rm;
RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K., Sasaki T.;
RT "Sequence of human parvovirus B19 isolates from patients with rheumatoid arthritis.";
Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AB030694; BAA90293.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat.1.
SQ SEQUENCE 554 AA; 60839 MW; 0DB958B33C73EF64 CRC64;

Query Match
Best Local Similarity 100.0%; Score 110; DB 12; Length 554;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 391 TYPFNKGTQOYTDQIERPLM 410

RESULT 25
ID Q9WKL9 PRELIMINARY; PRT; 554 AA.
AC Q9WKL9;
DT 01-NOV-1999 (TReMBLrel. 12, Created)

```

```

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Structural protein VP2.
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20014169; PubMed=10548133;
RA Hemaier A., Beckenlehner K., Wolf H., Lang B., Mocrow S.;
RT "Acute parvovirus B19 infection in connection with a flare of systemic lupus erythematosus in a female patient.";
DR J. Clin. Virol. 14:73-77(1999).
DR EMBL; AF13323; AAC99439.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat.1.
SQ SEQUENCE 554 AA; 60913 MW; 98FB588FF205C66F CRC64;

Query Match
Best Local Similarity 100.0%; Score 110; DB 12; Length 554;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 391 TYPFNKGTQOYTDQIERPLM 410

RESULT 26
ID Q8JNS4 PRELIMINARY; PRT; 554 AA.
AC Q8JNS4;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D91.1;
RX MEDLINE=22174902; PubMed=12186896;
RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G., Merlet J.F., Garbarg-Chenon A.;
RT "Genetic Diversity within Human Erythroviruses: Identification of Three Genotypes.";
DR J. Virol. 76:9124-9134(2002).
DR EMBL; AY083234; AAL91014.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat.1.
SQ SEQUENCE 554 AA; 60830 MW; A298575C083F24C8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 110; DB 12; Length 554;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPFNKGTQOYTDQIERPLM 20
DB 391 TYPFNKGTQOYTDQIERPLM 410

RESULT 27
ID Q9PZT8 PRELIMINARY; PRT; 760 AA.
AC Q9PZT8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE VP1/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kat12;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue."
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161224; AAD45912.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 760
SQ SEQUENCE 760 AA; 83403 MW; EA6E0145E3A0E5A CRC64;

Query Match 100.0%; Score 110; DB 12; Length 760;
Best Local Similarity 100.0%; Pred. No. 5.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYFPNKGTOOYTDQIERPLM 20
|||||
618 TYFPNKGTOOYTDQIERPLM 637

RESULT 28
Q9PZU0 PRELIMINARY; PRT; 761 AA.
AC Q9PZU0;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kat11;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue."
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161223; AAD45910.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 761
SQ SEQUENCE 761 AA; 83574 MW; 73A84C930D473530 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 761;
Best Local Similarity 100.0%; Pred. No. 5.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIERPLM 20
|||||
Db 618 TYFPNKGTOOYTDQIERPLM 637

RESULT 29
Q9PZT6 PRELIMINARY; PRT; 765 AA.
AC Q9PZT6;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kat13;

RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue."
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161225; AAD45915.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 765
SQ SEQUENCE 765 AA; 83998 MW; 89E2546086DCDBE8 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 765;
Best Local Similarity 100.0%; Pred. No. 5.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIERPLM 20
|||||
Db 618 TYFPNKGTOOYTDQIERPLM 637

RESULT 30
Q9PZT4 PRELIMINARY; PRT; 769 AA.
AC Q9PZT4;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kat14;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue."
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161226; AAD45917.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 769
SQ SEQUENCE 769 AA; 84578 MW; 0749D46E5CA7BB68 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 769;
Best Local Similarity 100.0%; Pred. No. 5.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIERPLM 20
|||||
Db 618 TYFPNKGTOOYTDQIERPLM 637

RESULT 31
Q913X1 PRELIMINARY; PRT; 773 AA.
AC Q913X1;
DT 01-DEC-2001 (TRENBLREL. 19, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Minor virus capsid protein VPI (Fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lali;
RA Hokynar K., Soderlund-Venemo M., Ranki A., Kiviluoto O., Partio E.K.,
Hedman K.;

RT "A new parvovirus B19 genotype persistent in skin."
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY044266; AAK95571.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 773
SQ SEQUENCE 773 AA; 85164 MW; D60CCA6F90B05378 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 5.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIERPLM 20
Db 618 TYFPNKGTOQYTDQIERPLM 637

RESULT 32
Q913W8 PRELIMINARY; PRT; 773 AA.

Q913W8;
01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Minor virus capsid protein VP1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

CK NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.

RA STRAIN=HAM;
RC Hokynar K., Soderlund-Venermo M., Ranki A., Kiviluoto O., Partio E.K.,

RA "A new parvovirus B19 genotype persistent in skin."
RT Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.

RL EMBL: AY044268; AAK95574.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 773
SQ SEQUENCE 773 AA; 85140 MW; 2B86F3245F9CE50 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 5.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIERPLM 20
618 TYFPNKGTOQYTDQIERPLM 637

RESULT 33

Q8JN56 PRELIMINARY; PRT; 781 AA.

AC Q8JN56;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE VP1.

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
CK NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=D91.1;
RX MEDLINE=22174902; PubMed=12186896;
RA Servais A., Laperche S., Lallemand F., Marinho V., De Saint Maur G.,

RA Merlet J.F., Garbarg-Chenon A.;
RT "Genetic diversity within Human Erythroviruses: Identification of
RT Three Genotypes";

RL J. Virol. 76:9124-9134(2002).
DR EMBL: AY083234; AAL91013.1; -
DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86142 MW; 4EB71FEDD41FC8F3 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 5.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIERPLM 20
Db 618 TYFPNKGTOQYTDQIERPLM 637

RESULT 34
Q65789 PRELIMINARY; PRT; 781 AA.

AC Q65789;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VP1.

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
CK NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=VEN1;
RA Echevarria Mayo J.E., Erdman D.D.;

RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL: U31358; AAA83558.1; -
DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 85985 MW; E1D5FDE230935DD CRC64;

Query Match 100.0%; Score 110; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 5.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIERPLM 20
Db 618 TYFPNKGTOQYTDQIERPLM 637

RESULT 35
P90223 PRELIMINARY; PRT; 781 AA.

AC P90223;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
CK NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=USA2;
RX MEDLINE=97081188; PubMed=8922470;
RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;

RT "Genetic diversity of human parvovirus B19: sequence analysis of the
RT VP1/VP2 gene from multiple isolates";

RL J. Gen. Virol. 77:2767-2774(1996).
DR EMBL: U38514; AAB47796.1; -
DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86070 MW; C1D911E39CF04A88 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 5.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIERPLM 20
Db 618 TYFPNKGTOQYTDQIERPLM 637

RESULT 36

Q85191 PRELIMINARY; PRT; 781 AA.
 AC Q85191;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Orf2 protein.
 GN ORF2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemmer A., Von Pöblich A., Giegler A., Casinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates.";
 EMBL; 270599; CAA94513.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86026 MW; AA02577B683BDB2D CRC64;

Query Match 100.0%; Score 110; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 5.3e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYPNKGTOOYTDQIERPLM 20
 |||||
 Db 618 TYPNKGTOOYTDQIERPLM 637

RESULT 37

P90221 PRELIMINARY; PRT; 781 AA.
 AC P90221;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=KOR2;
 MEDLINE=97081188; PubMed=8922470;
 Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the
 RT VP1/VP2 gene from multiple isolates.";
 RL J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38511; AAB47793.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86070 MW; 9FA830083F6F1357 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 5.3e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYPNKGTOOYTDQIERPLM 20
 |||||
 Db 618 TYPNKGTOOYTDQIERPLM 637

RESULT 38

P89318 PRELIMINARY; PRT; 781 AA.
 AC P89318;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=JAP1;
 MEDLINE=97081188; PubMed=8922470;
 Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the
 RT VP1/VP2 gene from multiple isolates.";
 RL J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38509; AAB47791.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86064 MW; 3AE65CE69096339 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 5.3e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYPNKGTOOYTDQIERPLM 20
 |||||
 Db 618 TYPNKGTOOYTDQIERPLM 637

RESULT 39

Q8UYE3 PRELIMINARY; PRT; 781 AA.
 AC Q8UYE3;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE 7.5 kDa protein.
 OS Erythrovirus A6.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=182494;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen Q.T., Wong S., Brown K.B.;
 RT "Identification and characterization of a second novel human
 RT erythrovirus variant, A6.";
 RL Virology 0:0-0(2002).
 DR EMBL; AY064475; AAL55418.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86277 MW; E73AB63E4551128 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 5.3e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYPNKGTOOYTDQIERPLM 20
 |||||
 Db 618 TYPNKGTOOYTDQIERPLM 637

RESULT 40

P90224 PRELIMINARY; PRT; 781 AA.
 AC P90224;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=BR21;
 MEDLINE=97081188; PubMed=8922470;

RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the
 RT VP1/VP2 gene from multiple isolates."
 RL J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL: U38546; AAB47801.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86039 MW; C3A29EB1DB8AD378 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 5.3e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIERPLM 20
 |||||
 Db 618 TYFPNKGTOQYTDQIERPLM 637

Search completed: August 20, 2003, 09:24:03
 Time : 58.143 secs

1

THIS PAGE BLANK (USPTO)